Advances in PET and multimodality imaging

With emphasis on cancer of the head and neck and liver metastases from colorectal cancer

Wouter Vogel

Cover illustration

The recent advances in PET imaging are represented by FDG-PET images from patients in subsequent stages of maturation. From left to right: aged 3 months, 4 years, 12 years, and 25 years. The reflections show integration of the PET images with CT images, as the next step of development.

ISBN: 978-90-9022157-1

Production
Layout: W.V. Vogel, Amsterdam, the Netherlands
Printing: Gildeprint BV, Enschede, the Netherlands

Advances in PET and multimodality imaging

With emphasis on cancer of the head and neck and liver metastases from colorectal cancer

Een wetenschappelijke proeve
op het gebied van de Medische Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het College van Decanen
te verdedigen op maandag 19 november 2007
om 13:30 uur precies

door

Wouter V. Vogel
geboren op 16 juni 1973 te ‛s-Hertogenbosch

Promotores:	Prof. dr. W.J.G. Oyen
	Prof. dr. J.H.A.M. Kaanders
	Prof. dr. F.H.M. Corstens
Copromotor:	Dr. J.A. van Dalen
Manuscriptcommissie:	Prof. dr. J.W.H. Leer
	Prof. dr. A. Heerschap
	Prof. dr. R.A. Dierckx (UMCG)

Contents

Introduction

Diagnosis and therapy of cancer in the head and neck area

Diagnosis and therapy of liver metastases in colorectal cancer

Discussion and conclusions

1

General introduction and outlines

1.1 Advances in medical imaging

Diagnosis and therapy of disease have undergone significant improvement, especially in the field of medical imaging (1). In the last decades, large scale research has sparked impressive progression in both technical possibilities and clinical knowledge (2). Unfortunately, these two factors have not always evolved in parallel. The discrepancy between technical advances and supporting evidence has been particularly large in the field of medical imaging (3). Experimental imaging may be justifiable in a research setting, but large-scale implementation of new techniques in clinical practice needs to be supported by proper validation, and scientific evidence of the benefit. The status of recent advances in the fields of molecular and multimodality medical imaging, with an emphasis on cancer in the head and neck area and on liver metastases, is discussed in this thesis.

1.2 Imaging modalities

Imaging has taken a prominent place in the spectrum of medical diagnostic procedures. Aspects of the human body can be visualized using different imaging modalities, each with their respective advantages and disadvantages.

Anatomical imaging

Visualization of anatomy in vivo can be performed with techniques such as conventional X-ray, ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI). These modalities all provide tools for high-resolution localization and delineation of normal organs and tissues as well as disease localizations, thus contributing to many diagnostic procedures, and to selection and planning of therapeutic procedures. Nevertheless, recognition of disease may be difficult using merely anatomical imaging, due to limited discrimination of tissue types.

Functional imaging

Visualization of dynamical tissue characteristics in vivo is known as functional imaging. Anatomical imaging modalities can be adapted to provide information about tissue characteristics by implementation of a dynamical component, for example CT imaging with multiphase intravenous contrast to provide information about local tissue perfusion (4). Other examples include ultrasound using air bubbles (5) and MRI using gadolinium (6). A disadvantage of contrast-enhanced anatomical imaging lies in the relatively high dose of applied contrast needed to provide sufficient signal for imaging. This limits the technique to molecules that are tolerated in very high doses, and thus are not metabolically active.

Molecular imaging

When biologically relevant molecules are imaged in physiological quantities, the related metabolical pathways can be visualized, and tissues can be characterized in vivo (7), with clinical implications for many benign (8) and malignant diseases (9).

In traditional nuclear medicine, molecules are labeled with gamma emitting radioisotopes to allow dynamical quantitative visualization of the biodistribution in time and place. Over the last decade, clinical molecular imaging has expanded further with the use of positron emitting radionuclides (10). The favorable radioactive decay of a positron, providing two simultaneous annihilation photons in opposite directions, permits imaging using a positron emission tomography (PET) scanner with a relatively good spatial and temporal resolution, and with a superior sensitivity (11). The technique permits in-vivo imaging of many biologically active molecules, such as peptides, hormones, antibodies, and pharmaceuticals, in physiological quantities. This allows quantitative evaluation of a wide variety of tissue parameters, such as metabolism, proliferation, hypoxia and receptor expression. The typical example is quantitative visualization of glucose metabolism using the radiopharmaceutical ^{18}F-fluor-deoxy-glucose (FDG), as a sensitive tool for the detection of cancer (figure 1.1).

While the clinical indications for PET continue to increase, the technique is subject to further improvement in many areas (12). Examples are improvement of image quality, better selection of indications, introduction of new radiopharmaceuticals, and integration with anatomical imaging modalities.

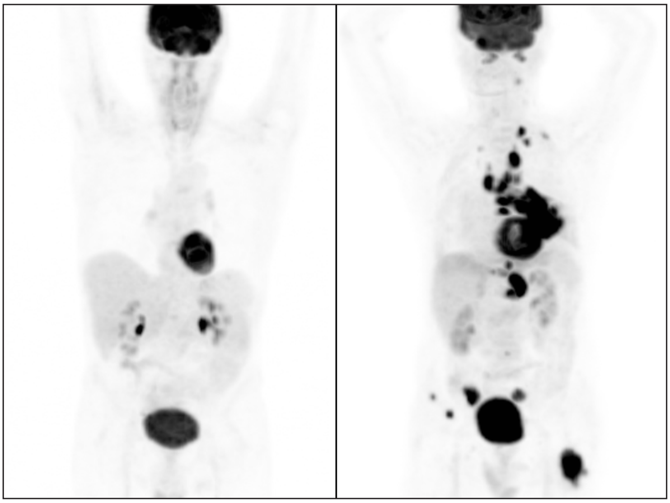


Figure 1.1
Visualization of cellular glucose metabolism with ^{18}F-fluor-deoxy-glucose in PET. Shown are maximum intensity projections from anterior from a normal healthy person (left), and a patient suffering from advanced stage lung cancer (right). The cancer lesions (in the left lung, lymph nodes, left adrenal, and multiple bones) show a high glucose metabolism, associated with tumor growth.

1.3 Multimodality imaging

Optimal selection of therapeutic strategies often requires a wide spectrum of clinical information, such as the activity and cellular characteristics of a disease, the number, size, and distribution of disease localizations, and involvement of adjacent structures.

The joined capabilities of anatomical imaging (high-resolution anatomical localization) and molecular imaging (sensitive qualitative and quantitative disease detection) have led to the practice of combining images, as illustrated in figure 1.2. This approach may provide synergistic diagnostic information (13). The correlated image sets may be acquired with separate imaging devices, or using an integrated scanning device. Both approaches are currently in use in clinical practice for the combination of PET and CT images, although both correlation strategies have their specific issues.

In chapter 2 the concept of multimodality PET/CT imaging is discussed, addressing both the implementations of software-fused separate image sets, and of an integrated hybrid scanning device.

The abovementioned molecular and combined imaging modalities are currently finding their applications in clinical diagnostic imaging. The advances in molecular imaging with PET and multimodality imaging with PET/CT, in malignancy in the head and neck, and to metastatic disease in the liver, are discussed in the following chapters.

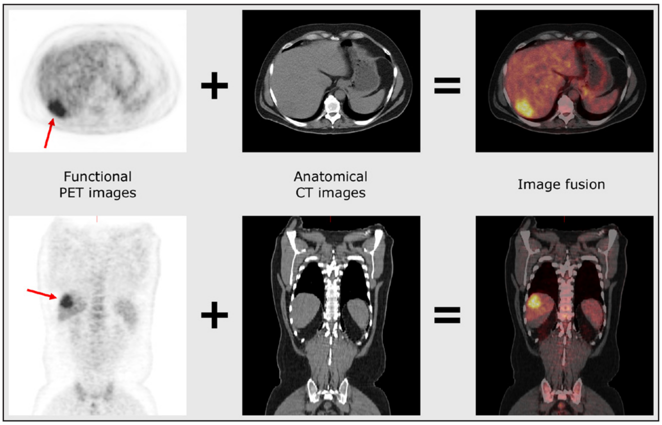


Figure 1.2
Illustration of integrated multi-modality imaging. Shown are transverse (top) and coronal slices (bottom) through the liver. A liver metastasis (arrow) is clearly visible on functional images from PET (A), and anatomy is well recognizable on CT images (B). Image fusion allows accurate localization of the metastasis within the liver (C).

1.4 Cancer of the head and neck area

Cancer of the head and neck area often poses complicated dilemmas (14). The primary tumor may compromise delicate structures adjacent to it, and the disease may spread rapidly and extensively throughout lymph nodes in the neck, and eventually throughout the body. Selection of an effective and yet sparing therapeutic strategy depends on accurate determination of the disease stage, which relies largely on imaging of the head and neck area, with a possible role for both anatomical and functional imaging (15). In the following chapters the advances of FDG-PET imaging in cancer of the head and neck area are evaluated.

Improving clinical imaging

In chapter 3 optimization of the image quality in FDG-PET imaging of the head and neck area is pursued. The parameters that influence image reconstruction from detected annihilation photons are discussed, and optimal settings are determined for the relatively low attenuation of photons in the head and neck area. The impact of these improvements on the diagnostic accuracy for detection of primary tumors and lymph node metastases is demonstrated.
In chapter 4 the clinical application of optimized FDG-PET imaging in oncology of the head and neck area is discussed. Currently, patients with no apparent lymph node metastases on conventional anatomical imaging undergo lymph node dissection for definitive staging. This invasive strategy has side-effects, and in retrospect often turns out to have been unnecessary. The study addresses the question whether the application of FDG-PET, with improved image quality, can improve the selection of patients for lymph node dissection.

Optimization of image fusion

Fusion of PET images with anatomical images may further improve diagnostic value, but only when performed adequately. This is of significant importance in imaging of the head and neck area, where structures are small and erroneous correlations are easily made. This implies a need for image sets with identical anatomical orientation, to allow reliable correlative reading.

In chapter 5 the presence of a difference in image size between PET and CT is discussed. Such differences may complicate image fusion. The impact of correction of such a difference on the accuracy of image registration is discussed.
In chapter 6 the process of anatomical image registration of PET and CT images is discussed. Several techniques are available for optimization of anatomical image registration, each with a different performance in the specific situation of the head and neck area. The available techniques are compared and their registration accuracies are determined, in order to validate the process for external beam radiation therapy planning.

Application in therapy planning

In chapter 7 the impact of multimodality PET/CT imaging on external beam radiation treatment planning is evaluated. Effective treatment, as well as sparing of normal tissues, depends on discrimination of normal and malignant tissues. Delineation of malignant tissue may be difficult on both anatomical and molecular images. Different techniques for delineation of a primary tumor in the head and neck area on integrated PET and CT images are evaluated, and an adequate strategy is suggested.

Improving specificity

In chapter 8 the application of ^{18}F-fluor-deoxy-thymidine (FLT) as an alternative to FDG for diagnostic PET imaging in the head and neck area is discussed. FDG suffers from a rather low specificity for imaging of lymph node metastases, due to the visualization of metabolic activity in non-malignant reactive lymph nodes. FLT, as a marker of proliferation rate, may be more specific. The value of FLT for the detection of lymph node metastases in the head and neck area is evaluated, to determine the applicability for diagnosis and external beam radiation treatment planning.

1.5 Liver metastases in colorectal cancer

Liver metastases from colorectal carcinoma remain a diagnostic challenge, despite ongoing advances in anatomical and molecular imaging methods (16). Locoregional therapeutic strategies could benefit from more accurate detection, localization, and delineation of lesions in the liver. Combined anatomical and functional imaging with CT and PET may prove valuable, but anatomical registration of the image sets is complicated in the upper abdominal area due to breathing (17). This applies to the correlation of separate PET and CT image sets, as well as integrated image sets from hybrid PET/CT.

PET/CT image fusion

In chapter 9 the specific benefits and problems of combined PET/CT imaging in the detection of recurrent colorectal carcinoma, and liver metastases in particular, are discussed.

In chapter 10 the approach of multimodality imaging with software-based image registration of stand-alone FDG-PET and CT images is evaluated. A novel organ-focused method for anatomical image registration of the liver is introduced and compared with several available methods. The accuracy in image registration is determined, in order to validate the process for diagnostic imaging of the liver.

In chapter 11 the approach of multimodality imaging with a hybrid PET/CT scanner is evaluated. This technique may be preferable over software-based image fusion for logistical reasons, but bears a risk for introduction of technique-related issues. The extent and impact of these issues are analyzed, and are compared with software-based image fusion. Recommendations are formulated to minimize these problems, for optimal imaging of the liver.

Tumor localization

Novel locoregional therapies of liver metastases, such as partial liver resection or radiofrequency ablation (RFA) of individual lesions, depends on accurate localization and delineation of lesions in relation to anatomical structures such as blood vessels.

In chapter 12, a novel algorithm for delineation of liver metastases on FDG-PET is presented. A liver metastasis appears on PET as a blurry hotspot, of which the apparent size depends on both the intensity of FDG-uptake and the real size of the lesion. An iterative thresholding technique is introduced, and compared with other available methods for tumor demarcation on PET.

In chapter 13 this novel demarcation method is applied to a series of clinical images, and the results are compared with conventional delineation techniques on PET and CT, with evaluation at pathology as a gold standard. An optimal strategy for delineation and measurement of liver metastases from colorectal cancer using FDG-PET is derived, and is validated against the CT imaging and histopathological examination.

1.6 References

1. Doi K. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. Phys. Med. Biol. 2006;51: R5-27
2. Looking back on the millennium in medicine. N. Engl. J. Med. 2000;342:42-49
3. Iglehart JK. The new era of medical imaging - progress and pitfalls. N. Engl. J. Med. 2006;354:2822-2828
4. Baron RL. Understanding and optimizing use of contrast material for CT of the liver. AJR Am. J. Roentgenol. 1994;163:323-331
5. Calliada F, Campani R, Bottinelli O, Bozzini A, and Sommaruga MG. Ultrasound contrast agents: basic principles. Eur. J. Radiol. 1998;27 Suppl 2:S157-S160
6. de Roos A, Doornbos J, Baleriaux D, Bloem HL, and Falke TH. Clinical applications of gadolinium-DTPA in MRI. Magn Reson. Annu. 1988;113-145
7. Jaffer FA and Weissleder R. Molecular imaging in the clinical arena. JAMA. 2005;293:855-862

8. Vos FJ, Bleeker-Rovers CP, Corstens FH, Kullberg BJ, and Oyen WJ. FDG-PET for imaging of non-osseous infection and inflammation. Q. J. Nucl. Med. Mol. Imaging. 2006;50:121-130

9. Weissleder R. Molecular imaging in cancer. Science. 2006;312:1168-1171

10. Bomanji JB, Costa DC, and Ell PJ. Clinical role of positron emission tomography in oncology. Lancet Oncol. 2001;2:157-164

11. Fischer BM, Olsen MW, Ley CD, Klausen TL, Mortensen J, Hojgaard L, and Kristjansen PE. How few cancer cells can be detected by positron emission tomography? A frequent question addressed by an in vitro study. Eur. J. Nucl. Med. Mol. Imaging. 2006;33:697-702

12. Cherry SR. The 2006 Henry N. Wagner Lecture: Of mice and men (and positrons) - advances in PET imaging technology. J. Nucl. Med. 2006;47:1735-1745

13. von Schulthess GK, Steinert HC, and Hany TF. Integrated PET/CT: current applications and future directions. Radiology. 2006;238:405-422

14. Chin D, Boyle GM, Porceddu S, Theile DR, Parsons PG, and Coman WB. Head and neck cancer: past, present and future. Expert. Rev. Anticancer Ther. 2006;6:1111-1118

15. Gibson MK and Forastiere AA. Multidisciplinary approaches in the management of advanced head and neck tumors: state of the art. Curr. Opin. Oncol. 2004;16:220-224

16. Bipat S, van Leeuwen MS, Comans EF, Pijl ME, Bossuyt PM, Zwinderman AH, and Stoker J. Colorectal liver metastases: CT, MR imaging, and PET for diagnosis--meta-analysis. Radiology. 2005;237:123-131

17. Nakamoto Y, Tatsumi M, Cohade C, Osman M, Marshall LT, and Wahl RL. Accuracy of image fusion of normal upper abdominal organs visualized with PET/CT. Eur. J. Nucl. Med. Mol. Imaging. 2003;30:597-602

2

PET/CT: Panacea, redundancy, or something in between?

Wouter V. Vogel [1]
Wim J.G. Oyen [1]
Jelle O. Barentsz [2]
Johannes H.A.M. Kaanders [3]
Frans H.M. Corstens [1]

Departments of Nuclear medicine [1], Radiology [2], and Radiation oncology [3],
Radboud University Nijmegen Medical Centre, the Netherlands

Journal of Nuclear Medicine, 2004
Volume 45, pages 15S-24S.

Abstract

In the past decade, the integration of anatomic imaging and functional imaging has emerged as a new and promising diagnostic tool. Developments in software have provided methods to integrate various modalities, such as PET, CT, MRI, and MR spectroscopy. The introduction of combined PET/CT scanners has boosted image fusion in this specific field and raised high expectations.

Image fusion can be performed at 3 different levels: visual fusion, software fusion, and hardware fusion, each having strengths, weaknesses, and issues inherent to technique. Visual fusion is the traditional side-by-side reviewing of 2 separate modalities. Software image fusion provides evaluation of 2 modalities in 1 integrated image set. True hardware fusion of PET and CT does not exist at present. Currently, hardware fusion refers to a PET/CT scanner that consists of separate scanners, which are positioned in line at a fixed distance, with automatic projection of the PET image over the CT image.

The suggested superiority of hardware fusion with these so-called hybrid PET/CT scanners over software fusion has sparked debate. Because scientific data that unequivocally show that state-of the-art software fusion is less accurate than hardware fusion (as provided in hybrid PET/CT scanners) are unavailable, the primacy of a combined PET/CT scanner over stand-alone PET and CT is more a matter of belief than of science.

Further research comparing the overall performance of PET/CT scanners with that of separate scanners with software for image fusion is much needed. The continuous development of better software for image fusion and respiratory and cardiac gating is also needed, not only for PET and CT imaging but also for fusion of PET with MRI and CT with MRI.

2.1 Introduction

In the past decade, functional imaging with ^{18}F-FDG PET has been the fastest growing diagnostic modality in oncology. The high sensitivity for depicting increased metabolism in a wide variety of malignancies adds significant accuracy to many diagnostic regimens compared with anatomic imaging only (CT, MRI, ultrasound). For several reasons, anatomical and functional imaging have been integrated into one diagnostic modality that is known as image fusion.

Image fusion can be performed at 3 different levels: visual fusion, software fusion, and hardware fusion. In traditional visual image fusion, the physician compares 2 separate imaging modalities viewed next to each other. The fusion takes place in his or her mind. In soft- and hardware image fusion, the results of both procedures are overlaid in an integrated set of images. It is the suggested superiority of hardware fusion in hybrid PET/CT scanners over software fusion that has sparked current discussions.

Software for image fusion has been developed by various vendors and is universally applicable to all sorts of image sets. True hardware fusion of PET and CT does not exist at present. It would require the use of a single detector system that registers 2 image sets at the same time (e.g., 511-keV γ-rays from ^{18}F-FDG and x-rays from CT). An alternative solution is a combined device with separate CT and PET scanners positioned in line. Several companies adopted this principle, and so-called hybrid PET/CT scanners are now widely available commercially.

Although hybrid PET/CT scanners are advertised extensively as the latest achievement in modern technology and as "state-of-the-art" and "must-haves", independent research on real benefits has only just begun. At this moment, few objective results have been produced that show the necessity of a combined PET/CT scanner or its advantages over software fusion. The current debate is mainly led by commercial companies and individual physicians' expert opinions. Therefore, comments and concerns are justified. Objective and independent grounds and arguments are discussed here, both for and against the use of image fusion in the modern daily clinical practice of diagnostic imaging and including both soft- and hardware fusion.

Anatomical imaging

Imaging modalities such as CT and ultrasound adequately reflect normal anatomy and anatomical changes. MRI has similar or even better potential for depicting anatomy and also provides additional tools for functional imaging. In recent years, the spatial resolution of these techniques has improved greatly, now significantly <1 mm. This permits accurate recognition and delineation of organs and structures, especially with the use of contrast-enhancing agents. These procedures are widely available and hold a solid position in the diagnostic algorithm of many disease entities. Despite these advances, anatomical imaging may be unable to differentiate between normal and pathological tissues with similar densities. It provides relatively little information about the viability or metabolic activity of organs and lesions, thus lacking sufficient sensitivity and specificity to answer a number of important clinical questions. Well-known problems are the differentiation of viable tumor from scar tissue after external beam radiation or chemotherapy, the detection of isodense metastases in the liver (in particular in a deformed liver after surgical procedures or in liver cirrhosis), and the detection of metastases in normal-size lymph nodes.

Functional imaging

Functional imaging with ^{18}F-FDG PET scanning permits differentiation of viable malignant tissue or active infection from normal tissue and from nonviable remnants by direct visualization of metabolic activity in vivo. Other tracers currently under development may prove useful in visualizing other important parameters, such as DNA synthesis, mitotic activity, protein synthesis, local ischemia, and expression of tumor-specific receptors. Despite high contrast resolution, major drawbacks of PET scanning are the relatively low spatial resolution of images (at present in the range of 4–6 mm and physically limited to about 1-2 mm) and poor recognition and delineation of anatomical structures. This may result in uncertainty or even failure in correctly localizing detected abnormalities. Recognized examples are lesions in the upper abdomen, situated near the border of organs, or between adjacent organs. Advanced MR techniques such as dynamic MRI and MR spectroscopy (MRS) are now available. These techniques also show functional aspects, such as vascularity, blood perfusion, oxygenation, and biochemical information. Moreover, MR contrast agents have been developed to label specific tissues. For example, ultrasmall superparamagnetic iron oxide particles specifically visualize macrophages and enable the differentiation of normal lymph node tissue from metastases (1).

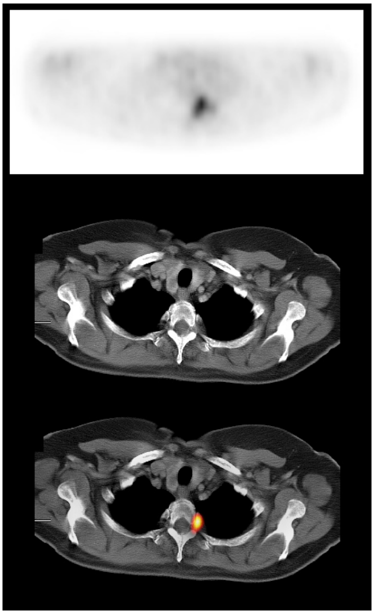


Figure 2.1
Example of software fusion of ^{18}F-FDG PET and CT images for lesion localization. (A) ^{18}F-FDG PET clearly showed a pathological lesion somewhere in the upper thoracic aperture. More precise localization was not possible due to lack of anatomic information. (B) The lesion was not found retrospectively on diagnostic CT images using visual fusion. (C) Software image fusion localized the lesion in a rib near the costovertebral joint.

2.2 Benefits of image fusion

The limitations in separate CT and PET imaging may be compensated for when the 2 modalities are used in a complementary way. High-resolution anatomical information produced by CT adds significant information to tissue characterization delivered by PET. In addition, fusion of high-resolution MRI anatomical and functional information with PET will provide an extra dimension. When applying the integration of different imaging modalities, image fusion becomes an issue. Adequate anatomical alignment of both image sets permits convenient visualization of all information in one study.

Diagnostic effect

Improved lesion characterization and localization will result in increased diagnostic accuracy, which is recognized as a beneficial diagnostic effect. However, better accuracy in staging and restaging of disease is only relevant when it leads to changes in patient management (e.g., by decreasing the number of indications for invasive procedures). Also, improved lesion localization may lead to better results in other successive diagnostic procedures (e.g., easier CT-guided biopsy). Figure 2.1 is an example of improved lesion localization with software image fusion.

Integration of multiple functional imaging techniques, such as PET studies with various tracers or functional MRI with contrast enhancement and MRS, results in comprehensive in vivo tissue characterization.

Effect on therapy

PET images can be implemented in radiotherapy treatment planning and may be of particular value for high-precision techniques, such as intensity-modulated radiotherapy (IMRT). With IMRT, different dose prescriptions can be delivered to multiple target sites with

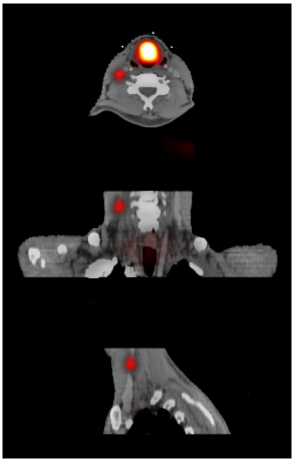


Figure 2.2
Example of software fusion of ^{18}F-FDG PET and CT images of primary laryngeal carcinoma for IMRT field planning. Patient was scanned for both ^{18}F-FDG PET and CT on a flat bed, in a personalized rigid radiotherapy mask covering head and shoulders to prevent positioning differences. A lymph node in neck that was only marginally enlarged on CT proved pathologic on ^{18}F-FDG PET and was included in the radiotherapy field. (A) Transverse slice through the pathologic lymph node. (B) Coronal slice. (C) Sagittal slice.

extremely high dose gradients between tumor and normal tissues. This places demands on the ability of conventional imaging techniques to localize and delineate tumors. PET scanning can detect additional lesions or may provide complementary information to facilitate the interpretation of equivocal CT findings (e.g., marginally enlarged lymph nodes). Consequent adjustments of the radiotherapy target volume will have a direct effect on the chances of cure, and on the risk and level of side effects and on complications.

In addition, functional PET imaging may identify tumors or regions within tumors with increased radioresistance. Examples are tumor hypoxia and areas of very active tumor cell proliferation that can be detected by specific tracers (2). A next step in the development of IMRT will be the integration of anatomical and functional information into a biological target volume (3). Using the ability of IMRT to deliver nonuniform dose patterns, biological dose conformality can be pursued, creating higher doses to areas of increased radioresistance, and lower doses in areas of high radiosensitivity. Figures 2.2 and 2.3 demonstrate the use of software image fusion in IMRT planning. Figure 2.4 is an example of a resulting IMRT field. MRI and dynamic and spectroscopic MR information also can be used to target IMRT and have been shown to very accurately localize prostate cancer (4,5). Fusion of this functional MR information with CT has also been shown to be possible (6).

The potential value of combined modalities is easy to recognize. Joined information may prove synergistic. It allows accurate differentiation between pathological and normal tissue, with excellent resolution and localization, and it creates opportunities for further optimization of treatment. The question that remains to be answered is whether only a combined PET/CT scanner can live up to this expectation.

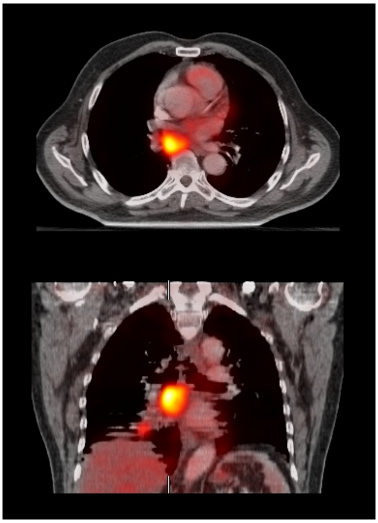


Figure 2.3
Example of software fusion of ^{18}F-FDG PET and CT images of primary non–small cell lung cancer with a lymph node metastasis in mediastinum for IMRT field planning. The patient was scanned for both ^{18}F-FDG PET and CT on a flat scanning bed, with the arms up in a rigid customized support system to minimize positioning differences. CT images were acquired during free tidal breathing, accounting for huge artifacts. ^{18}F-FDG PET images were also acquired during free tidal breathing but show much more reliable delineation of tumor tissue. (A) Transverse slice. (B) Coronal slice.

2.3 Levels of image fusion

When reading two image studies such as PET and CT, three levels of image fusion are possible: visual, software, and hardware fusion. When two studies need to be correlated, what level of image fusion is preferred or needed?

Often no need for image fusion

A PET scan without a CT scan is not useless. On the contrary, most PET scanners today operate as stand alones and perform quite well. In most cases, PET images contain enough information to answer relevant clinical questions. Additional information necessary to interpret the images should be retrieved from the referring physician or from previous imaging, such as CT, ultrasound, or MRI.

Visual fusion

The reviewer traditionally has a film print or digital display of a previously recorded CT scan next to PET images, and overlays the images in his or her mind while reviewing. This is called visual fusion. Based on clinical experience, Jager et al. (7) estimated that there is a need to look at CT images in only about 20% of cases. They also stated that, in the majority of these selected cases, visual fusion gives sufficient information with no actual need for or additional value from soft- or hardware image fusion. Others emphasize the necessity of an extensive review of CT images in all cases. These 2 opinions, neither founded on evidence-based data, serve as a source of controversy.

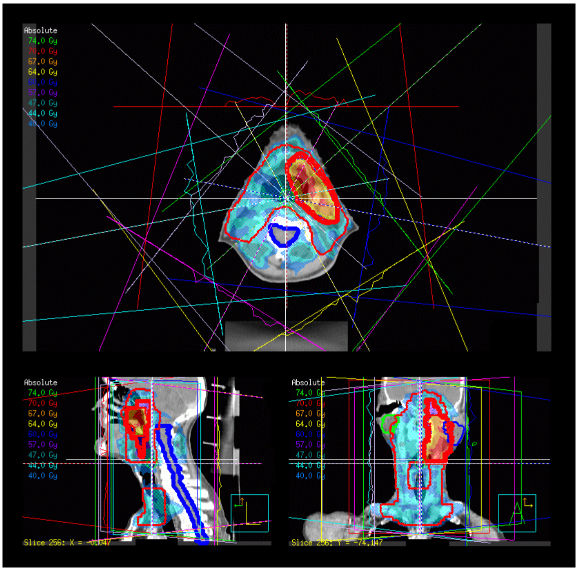


Figure 2.4
Sample IMRT field planning of oropharyngeal squamous cell carcinoma. Different levels of radiation dose are drawn in red in a 3-dimensional space. Painted in blue is the spinal cord that will be spared. Reverse planning was used to calculate the optimal shape and intensity of 7 beams.

Software fusion

Integration of separate PET and CT image sets into a single study can be achieved with software fusion. Several commercial packages have been developed for this purpose. These packages share a similar set of functions for fusion and visualization.

The most elementary method of image fusion is manual positioning of one scan relative to the other in a 3-dimensional space, with 9 free parameters for adjustment of position, size, and rotation in 3 directions each. This is a very time-consuming procedure and, more important, is highly operator dependent and not very accurate. Marking of recognizable points in both scans (so-called landmarks, either anatomic or artificial) was introduced initially for optimization, but lack of such anatomic points on PET and low correlation of external fiducial markers with internal organ positions limits the applicability of this technique. The true power of these software packages lies in the use of automatic optimization algorithms, which eliminate operator dependency and increase accuracy by considering all parts of both scans simultaneously. Numerous algorithms have been developed for this purpose. Examples include count difference, shape difference, mutual information, normalized mutual information, and others. All currently used algorithms produce fused images, but the results vary in accuracy. It is unclear which algorithm performs best in a specific situation. However, mutual information is generally accepted as the most robust procedure. Both PET emission and transmission images (or even a combination of those in a given ratio) may be considered when optimizing (8). Only rigid transformation algorithms are currently commercially available and widely used. In the near future, elastic transformation will become available, possibly allowing automatic correction of small breathing artefacts and

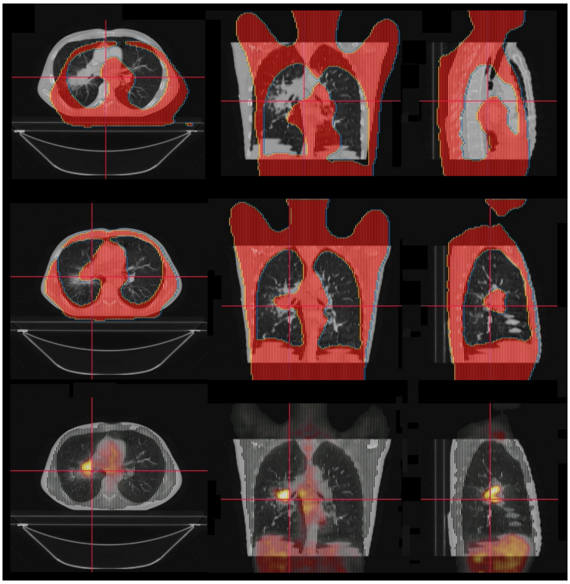


Figure 2.5
Example of software fusion procedure for ^{18}F-FDG PET and CT images. (A) After loading both scans in fusion software, a large shift exists. PET transmission images in CT are displayed for easier recognition of body outlines. (B) Optimized fusion after automatic registration with mutual information algorithm. (C) Final result with ^{18}F-FDG-PET emission images displayed in CT.

positioning differences (9,10). Still, it is unlikely that advanced techniques in software fusion will ever be able to correct for extreme positioning differences and motion artifacts. Furthermore, patient logistics and network connectivity are relevant issues when performing software image fusion, because each may lead to additional delay.

The software fusion process typically consists of several steps. First, 2 image sets must be loaded into the computer system running the fusion application. If all geometric parameters are correctly defined, both image sets will appear in an integrated image set with a certain displacement and possibly a small rotation along 3 axes. A quick correction of the most obvious displacement may be performed manually. Additional optimization must be achieved by automatic registration, which also will be fast. The resulting integrated image set may be viewed or saved to a picture archiving and communication system. Figure 2.5 shows an example of the procedure.

Hardware fusion

As stated previously, true hardware image fusion of PET and CT is not yet possible. The term hardware fusion currently refers to a PET/CT scanner, which includes separate scanners positioned in line at a fixed distance. In fact, current combined PET/CT scanners can be described as expensive patient positioning systems that facilitate obtaining a dedicated CT and a dedicated PET in one session with minimal patient movement, at the cost of occupation of the CT while scanning PET and vice versa. The acquired image sets are calibrated to be overlaid correctly within a certain error but are not routinely corrected for breathing artefacts or accidental positioning changes between the 2 scans. Because of these inaccuracies, CT-based attenuation correction for the PET images, advocated for its speed of acquisition, will introduce certain artefacts. Software for optimization or correction of these errors in hybrid PET/CT images is not used currently.

Nevertheless, the use of a combined PET/CT scanner has several definite advantages over separate scanners. The patient lies on the same bed in the same position. Therefore, the image fusion procedure logistics are easier and probably less prone to artefacts resulting from positioning differences. In addition, total scanning time is shorter when using the CT for attenuation correction of PET data.

A point of controversy

The required level of image fusion remains open for debate. It is unclear whether the increased scanning speed and somewhat easier image fusion (when needed) are sufficient compensation for the higher radiation exposure and high cost of a combined PET/CT scanner. No studies have yet addressed this problem as a whole.

Lardinois et al. (11) studied hardware image fusion compared with CT alone, PET alone, and visual correlation, with histopathology as a gold standard in staging 49 patients with non–small cell lung cancer. Software image fusion was not performed in this study. The authors found integrated PET/CT significantly more accurate in staging the primary tumor than CT alone, PET alone, or visual image fusion. They reported improved accuracy in staging lymph nodes compared with PET alone, but no significant difference in accuracy was found between integrated PET/CT and visually correlated images. Integrated PET/CT provided additional information over visual fusion in 41% of cases, including more exact localization of tumor sites, precise evaluation of chest wall or mediastinal involvement, and differentiation between tumor and peritumoral inflammation or atelectasis. However, the key questions, how this related to the performance of software fusion and whether and to what extent patient management or therapy outcomes were changed, were not addressed.

2.4 Problems and limitations

When using image fusion in daily clinical practice, a certain level of accuracy in positioning one image set relative to the other must be achieved or, better, guaranteed. Does hardware image fusion perform better than software fusion?

Accuracy

The maximum achievable accuracy in image fusion depends on several factors. Most important are patient positioning differences, internal organ movements (e.g., breathing and motion of the heart), attenuation correction artefacts, and errors in the fusion procedure. Figure 2.6 shows the potential results of some of these effects. To a certain degree, these factors (except

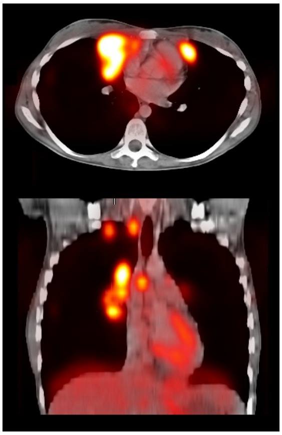


Figure 2.6
Example of failed software image fusion in a patient with B-cell lymphoma. CT images were acquired during breathhold at deep inspiration with the arms up. ^{18}F-FDG PET images were acquired during free tidal breathing with the arms down. The fusion images show major inaccuracies in upper thorax aperture, where pathological lymph nodes are mispositioned in the lung, and in region of diaphragm, which is misaligned. (A) Transverse slice. (B) Coronal slice.

for CT-based attenuation correction artefacts) are present in both soft- and hardware fusion. The presence of these errors must be checked and corrected when possible. This can be done with image registration software.

Patient positioning and movement

Patient positioning should be identical during the PET and CT scans. This implies scanning with the arms either up or down in both scans, with identical support material under the head and neck and with similarly shaped scanning beds. The choice of arm position depends on the purpose of the scan. In most diagnostic procedures, CT scans are acquired with the arms up. PET scans are usually acquired with the arms down for patient comfort during the relatively long scanning period. Compromises may be unavoidable in image fusion.

Patient movement should be minimized during and between scans by instructions and additional support material. Various types of custom-molded support cushions are available for this purpose. When very high accuracy is needed (e.g., in radiotherapy of the head and neck) customized head support devices and immobilizing masks are recommended. Despite best efforts, patient movement can never be ruled out as a source of error.

Patient positioning and motion errors are possible when imaging with separate scanners as well as with integrated PET/CT scanners, although the chance of large positioning differences between scans seems significantly lower in the latter. This problem requires more attention when using separate scanners but can be largely overcome when handled appropriately. Equipment suppliers should facilitate this by providing the same bed for PET and CT.

Organ movement

Because of the duration of the scan, PET is acquired during free tidal breathing. This implies continuous movement of the thoracic wall, lungs, mediastinum, diaphragm, and upper abdominal organs, leading to displacement (up to several centimeters), rotation, and deformation of organs. Therefore, diagnostic CT scans are usually acquired during deep inspiration breath hold. Standard PET and CT, then, will inevitably contain major differences. The only available solution is the acquisition of the CT scan during free breathing (with the introduction of serious artefacts on CT) or during breath hold in a fixed position. Goerres et al. (12) reported the best match of CT and PET when scanning CT in unforced expiration but also recognized the inability of most patients to hold this position during a whole-body CT. A compromise suggested by Beyer et al. (13) allowed the patient to breathe free at the beginning and end of the CT scan and advocated unforced expiration during scanning of the lower lung and upper abdominal region. This procedure provides an acceptable although still not perfect result in image fusion. Reliability in the diaphragmatic region remains poor. Novel respiratory-gated PET scanning techniques are being developed but are not yet clinically used (14). This problem is not solved, and it persists in both integrated PET/CT and separate scanners.

Considerable motion and displacement of internal organs are often problematic in scans of the pelvic region. This is mainly caused by variations in bladder and rectal filling, which in turn can affect prostate and uterus position. This problem is probably best solved by keeping the interval between the scans as short as possible.

Although combined PET/CT scanning delivers readily aligned image sets, it is not perfect by default. The scans may be slightly misaligned as a result of a calibration error in the relative CT and PET scanner positions. Minor differences in patient positioning may occur, and differences in diaphragm position will exist. A minimal error of several millimeters is considered unavoidable. However, variations in positioning and initial misalignment will be larger when using separate scanners.

Scanner resolution and inaccuracy

The minimal error that can be visualized and corrected, either manually or with automatic registration, is limited by the resolution of both scans. For PET and CT this implies an error of approximately 5 mm in the transverse plane and at least 5 mm along the longitudinal axis, depending on the thickness of CT slices. Automatic registration may achieve a somewhat higher accuracy by considering all slices at the same time, but this is difficult to prove.

The total inaccuracy of fusion consists of the added inaccuracies caused by positioning differences, motion artefacts, and the fusion procedure itself. As a rule of thumb, the total minimal theoretic error can be estimated at approximately 10 mm in all directions and probably twice that in the diaphragmatic region. The minimal error may be less in rigid body parts, such as the brain. The minimal error is unavoidable in image fusion, whether a separate or combined PET/CT scanner is used. Many soft- and hardware solutions claim higher accuracy in image fusion, up to 1–2 mm in all directions. Such claims are based on phantom studies and disregard the specific problems inherent when scanning living subjects.

Acceptable inaccuracy

The required accuracy of image fusion depends on the clinical question that needs to be answered. To correctly distinguish 2 lesions, the error must be smaller than the size of the lesions. The smallest lesion that can be detected with PET is just below 1 cm in diameter. Therefore the maximum level of accuracy that can be achieved, and is required, is also in the range of 1 cm in all directions.

In radiotherapy of the head and neck area, with optimal positioning and immobilization, the error in dose delivery is in the order of 2 mm or less (15). Such degree of accuracy cannot be achieved with a PET scanner with 5-mm spatial resolution. The total error of PET/CT fusion must be considered when defining the target volume in radiotherapy. As a consequence, the target volume will increase. On the other hand, the interobserver variations in contouring tumor volumes can be significantly reduced because of the higher specificity of PET information (16). These factors must be weighed against the benefits of higher sensitivity of tumor detection and the availability of additional functional information. It opens a whole new field of research and development.

Because additional error has direct consequences for the size of radiation volumes, every effort must be made to achieve the highest possible accuracy. This means that all imaging, as well as the actual radiotherapy, must be performed with the patient locked in a rigid mask that prevents any movement or variation in positioning. In radiotherapy field planning of the lungs and mediastinum, the CT scan is recorded during free breathing to mimic the situation during radiotherapy. Under these conditions, the CT has an effective total error of several centimeters. Here, the PET scan is considered to have a better resolution, as illustrated in figure 2.3. New developments in radiotherapy treatment planning and delivery address this problem. The use of multiple, so-called "slow" CT scans, which use long sampling times for image registration, allows better capture of tumor movement and generates more reproducible target volumes (17). The precision of dose delivery can be improved with respiratory-gated radiotherapy, where dose delivery is adjusted to the breathing motion of organs. Again, for these issues, no significant difference exists between the use of 2 separate scanners and a combined PET/CT scanner.

All errors mentioned here are theoretic estimates. Unfortunately, the real error made in image fusion cannot be assessed objectively, because no gold standard is available. However, methods to evaluate the quality of image fusion have been developed and are used in research settings (18). The human eye is the only available instrument to recognize specific errors in individual cases. This means that for both soft- and hardware solutions, image fusion has no "guaranteed" accuracy—only estimated minimal errors.

2.5 Attenuation correction

Advantages

Most institutions perform a transmission scan to correct the emission scan for photon attenuation. The time needed for the transmission scan using radioactive sources is about 30%–40% of the total scanning time. The CT scan of a PET/CT study can also be used to correct the PET emission for attenuation, thus decreasing the total PET scanning time. It is obvious that such reduction in scanning time will not be achieved when 2-dimensional emission data alone are used to read the PET images.

Another advantage of CT-based attenuation correction is a lower noise level in the CT images compared with that in traditional transmission images, resulting in corrected PET images with less noise.

Disadvantages

Using CT data to correct PET emission data to reduce total scanning time is tempting. However, significant additional errors may be introduced. As indicated previously, positioning and breathing protocol differences between the CT and PET scans will result not only in image fusion errors but also will introduce quite serious artefacts in the attenuation-corrected PET images, especially in the diaphragm region where lesions may be projected in the wrong organ (19). The use of lower energy x-rays instead of a 511-keV transmission source also will result in more artefacts in the area of metallic implants. Overcorrection will result in false hotspots, unless the attenuation correction algorithm is specifically adapted to this problem (20). Serious distortion effects have been demonstrated on PET near dental implants and prostheses (21,22). Dizendorf et al. (23) reported that contrast fluids have some effect but that this did not lead to serious misinterpretations. Uncorrected PET images should be viewed when using CT-based attenuation correction, thereby avoiding serious misinterpretations. In theory, it is possible to use CT data for attenuation correction of PET data as acquired on a separate PET scanner. However, such a procedure is considered cumbersome because of logistic implications and the very high risk of artefacts.

2.6 Additional arguments

The clinical and technical arguments mentioned previously are perhaps the most important, but the discussion is certainly not limited to these. Several other factors with regard to PET/CT should be considered.

CT image quality

CT scans acquired with a combined PET/CT scanner may not be of the best diagnostic quality. Because the main purpose of the CT scan in this combination is coregistration and attenuation correction, a single-slice, low-energy CT scan with relatively low contrast and resolution and a relatively large slice thickness is sufficient. In some modern scanners a high-resolution, multislice diagnostic CT scanner is available, but in most cases the patient has already undergone a high-quality diagnostic CT scan before referral for PET. This limits the need for a second state-of-the-art CT scan.

Radiation dose

In PET/CT scanning, the patient undergoes a CT scan, regardless of whether a diagnostic CT scan is already available. This implies a second radiation dose. A standard PET emission scan delivers a radiation dose to the body of 5–10 mSv. A traditional germanium-based transmission scan delivers no significant dose to the patient. When using a modern multislice CT with diagnostic settings, the received dose may be up to 10 mSv for the head, 20 mSv for the chest, and 20 mSv for the abdomen. This would be a significant increase in total radiation to the patient for PET scanning. This extra radiation dose can be decreased by using lower-energy x-rays and thicker slices that are sufficient for CT-based attenuation correction and simple lesion localization in PET/CT scanning. Nevertheless, a total-body CT dose of 10 mSv added to the normal PET emission dose is very common. In nuclear medicine, a trade-off between scanning speed and radiation dose always exists but usually is not that significant.

It is also worth considering that a combined PET/CT scanner delivers a CT scan with every PET scan, needed or not. The CT scan then should be used for diagnosis. It would be poor radiation safety practice to acquire a wholebody CT scan for faster attenuation correction only and not even look at the inherent diagnostic information afterwards. Eventually, implementation of PET/CT as a common tool in diagnosis may lead to an adaptation of the diagnostic strategy. In certain situations, obtaining a diagnostic CT before PET/CT may not be necessary. Such a strategy is currently under investigation for the diagnostic work-up of patients with suspected lung cancer, and some institutions have already implemented this in routine clinical practice for this diagnosis.

Combination with other modalities

The combination of PET and CT may be attractive, but the integration of PET with imaging modalities other than CT can have additional important advantages. MRI is superior to CT in visualization and delineation of soft tissue anatomy. The combination of PET with imaging modalities that do not contribute to radiation dose, such as MRI, would be preferable in terms of radiation safety. The fact that PET/CT is the only available integrated combination right now does not make it a panacea. In the near future, image fusion of PET with other modalities is certain to become an issue, and it would be unwise to direct all attention, efforts, and available budget to a single option. This may even slow down development in other important areas of image fusion.

Economy

The purchase of a combined PET/CT scanner is quite costly and must be adequately justified. The difference in price between a PET/CT unit and a dedicated PET scanner is sufficiently large to warrant considering spending the budget on other, more urgently needed items. This becomes even more of an issue when taking into account the fact that the high-end CT included in a PET/CT scanner will be operated for only 1 min per patient, for a total of only 10–20 min per day. The combination of 2 advanced imaging devices in 1 machine at least doubles the chance of technical failure. Personnel dually trained in PET and CT techniques are needed to adequately operate such a machine. Such factors increase the total costs of ownership.

Competence issues

When a nuclear medicine department begins using a CT scanner, competence issues may arise with the radiology department. Are the CT images used for localization and attenuation correction only? Are nuclear medicine physicians allowed to review the CT and use it as a diagnostic tool as well? Can these findings end up in the PET report? Should a radiologist always look at the CT images? Can this be done separately, or is a joined image reading session mandatory? On the other hand, the radiotherapy department may want to use a PET/CT scanner to improve radiation treatment planning. Competence and responsibilities of nuclear medicine physicians, radiologists, and radiation oncologists need to be defined.

2.7 Conclusions

Without reservation, the answer to the question of whether we need CT in combination with PET is yes. However, questions about the required level of image fusion are far from settled. In the authors' opinion, side-by-side reading of PET and CT (meaning visual image fusion) is sufficient and adequate in many cases. Moreover, scientific data are unavailable to demonstrate unequivocally that state-of-the-art software fusion is less accurate than the hardware fusion provided in current PET/CT scanners. Suggesting the primacy of a combined PET/CT scanner over stand-alone PET and CT is more a matter of belief than of science.

The possession of a PET/CT scanner does not automatically upgrade an institution to the top of its class, and, perhaps more important, an institution with a state-of-the-art PET scanner, state-of-the-art CT scanner, and state-of-the-art software for image fusion is definitely not second class. High quality should be defined by state-of-the-art PET, CT, adequate software for image fusion, and knowledgeable multidisciplinary interpretation. Those institutions starting with clinical PET can perform very well with a stand-alone, dedicated PET scanner, in combination with software fusion when needed. The extra expense of a PET/CT scanner should be weighed against other, possibly more urgent, needs.

Additional research comparing the overall performance of PET/CT scanners with that of separate scanners with software for image fusion is much needed. Furthermore, the continuous development of software for image fusion and respiratory and cardiac gating is not only needed for combined PET and CT imaging but also for fusion of PET with MRI and CT with MRI. For these reasons, it is counterproductive when only PET/CT scanners receive the full attention of the nuclear medicine community.

2.8 References

1. Harisinghani MG, Barentsz JO, Hahn PF, et al. MR lymphangiography for detection of minimal nodal disease in patients with prostate cancer. Academic Radiology. 2002;9:S312-S313.
2. Chao KSC, Bosch WR, Mutic S, et al. A novel approach to overcome hypoxic tumor resistance: Cu-ATSM-guided intensity-modulated radiation therapy. International Journal of Radiation Oncology Biology Physics. 2001;49:1171-1182.
3. Ling CC, Humm J, Larson S, et al. Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. International Journal of Radiation Oncology Biology Physics. 2000;47:551-560.
4. Engelbrecht MR, Huisman HJ, Laheij RJF, et al. Discrimination of prostate cancer from normal peripheral zone and central gland tissue by using dynamic contrast-enhanced MR imaging. Radiology. 2003;229:248-254.
5. Futterer JJ, Scheenen T, Welmers A, De la Rosette JJ, Blickman JG, Barentsz JO. Dynamic contrast-enhanced MRI using both high spatial and high temporal resolution combined with 3D-MR spectroscopy in prostate cancer imaging. Radiology. 2002;225:628-628.
6. Welmers A, Futterer JJ, Van Lin E, Scheenen TW, Huisman HJ, Barentsz JO. 3D image fusion of CT with dynamic contrast enhanced MRI and MR spectroscopy of the prostate using gold makers as fiducials. Radiology. 2002;225:350-351.
7. Jager PL, Slart RHJA, Corstens F, Hoekstra O, Teule J, Oyen WJG. PET-CT: a matter of opinion? Eur J Nucl Med Mol Imaging. 2003;30:470-471.
8. Skalski J, Wahl RL, Meyer CR. Comparison of mutual information-based warping accuracy for fusing body CT and PET by 2 methods: CT mapped onto PET emission scan versus CT mapped onto PET transmission scan. Journal of Nuclear Medicine. 2002;43:1184-1187.
9. Little JA, Hill DLG, Hawkes DJ. Deformations incorporating rigid structures. Computer Vision and Image Understanding. 1997;66:223-232.
10. Slomka PJ, Dey D, Przetak C, Aladl UE, Baum RP. Automated 3-dimensional registration of stand-alone F-18-FDG whole-body PET with CT. Journal of Nuclear Medicine. 2003;44:1156-1167.
11. Lardinois D, Weder W, Hany TF, et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. New England Journal of Medicine. 2003;348:2500-2507.
12. Goerres GW, Burger C, Schwitter MR, Heidelberg TNH, Seifert B, von Schulthess GK. PET/C of the abdomen: optimizing the patient breathing pattern. European Radiology. 2003;13:734-739.
13. Beyer T, Antoch G, Blodgett T, Freudenberg LF, Akhurst T, Mueller S. Dual-modality PET/CT imaging: the effect of respiratory motion on combined image quality in clinical oncology. European Journal of Nuclear Medicine and Molecular Imaging. 2003;30:588-596.

14. Nehmeh SA, Ford E, Rosenzweig K, et al. Gated Positron Emission Tomography: A technique for reducing lung tumor motion effect. Journal of Nuclear Medicine. 2001;42:34P-34P.

15. van Lin ENJTh, van der Vight L, Huizenga H, Kaanders JHAM, Visser AG. Set-up improvement in head and neck radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support. Radiother Oncol. 2003;68:137-148.

16. Caldwell CB, Mah K, Ung YC, et al. Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: The impact of (18)FDG-hybrid PET fusion. International Journal of Radiation Oncology Biology Physics. 2001;51:923-931.

17. Lagerwaard FJ, de Koste JRV, Nijssen-Visser MRJ, et al. Multiple “slow” CT scans for incorporating lung tumor mobility in radiotherapy planning. International Journal of Radiation Oncology Biology Physics. 2001;51:932-937.

18. Klabbers BM, de Munck JC, Slotman BJ, et al. Matching PET and CT scans of the head and neck area: Development of method and validation. Medical Physics. 2002;29:2230-2238.

19. Osman MM, Cohade C, Nakamoto Y, Marshall LT, Leal JP, Wahl RL. Clinically significant inaccurate localization of lesions with PET/CT: Frequency in 300 patients. Journal of Nuclear Medicine. 2003;44:240-243.

20. Goerres GW, Burger C, Kamel E, et al. Respiration-induced attenuation artifact at PET/CT: Technical considerations. Radiology. 2003;226:906-910.

21. Kamel EM, Burger C, Buck A, von Schulthess GK, Goerres GW. Impact of metallic dental implants on CT-based attenuation correction in a combined PET/CT scanner. European Radiology. 2003;13:724-728.

22. Goerres GW, Ziegler SI, Burger C, von Schulthess TBGK, Buck A. Artifacts at PET and PET/CT caused by metallic hip prosthetic material. Radiology. 2003;226:577-584.

23. Dizendorf E, Hany TF, Buck A, von Schulthess GK, Burger C. Cause and magnitude of the error induced by oral CT contrast agent in CT-based attenuation correction of PET emission studies. J Nucl Med. 2003;44:732-738.

3

Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy

Wouter V. Vogel [1]
Bart M. Wensing [2]
Jorn A. van Dalen [1]
Paul F.M. Krabbe [3]
Frank J.A. van den Hoogen [2]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1], Otorhinolaryngology and head and neck surgery [2], and Medical technology assessment [3], Radboud University Nijmegen Medical Centre, the Netherlands

European Journal of Nuclear Medicine and Molecular Imaging, 2005
Volume 32, pages 1276–1282.

Abstract

Reconstruction parameters are an important factor in PET image quality. In the head and neck area, where the level of photon attenuation is relatively low, standard whole-body reconstruction (SWR) parameters may lead to suboptimal results. The purpose of this study was to evaluate the impact of optimised head and neck reconstruction (OHR) parameters on image quality and diagnostic accuracy, using pathology as the gold standard.

Methods: SWR parameters consisted of 2 iterations, 8 subsets and a 6-mm Gaussian filter. Predetermined OHR parameters were 4 iterations, 16 subsets and a 5-mm Gaussian filter, generating images with increased spatial and contrast resolution but also with increased noise. SWR- and OHR-based FDG-PET images of 28 patients with malignancies in the head and neck area were evaluated for primary tumor and pathological lymph nodes. Diagnostic accuracy was determined by histopathological verification after lymph node dissection.

Results: Using OHR, sensitivity for detection of a primary tumor increased from 92% to 100%. Eleven additional lymph nodes were visualised in eight patients, resulting in an increased sensitivity for lymph node metastases from 11% to 44%. Specificity decreased from 89% to 74% owing to visualisation of small reactive lymph nodes. In total, using OHR, FDG-PET diagnosis improved in six patients (21%) at the expense of three additional false positives for lymph node metastasis (11%). Primary tumor SUV_{max} increased by 42%, indicating enhanced contrast resolution.

Conclusion: Image reconstruction adapted to low photon attenuation in the head and neck area may improve image quality and the diagnostic value of FDG-PET, despite a slightly higher false positive rate attributable to the fact that visualisation of FDG accumulation in benign reactive lymph nodes is also enhanced.

3.1 Introduction

Primary tumors in the head and neck area are often small at the time of discovery, especially when located in the oral cavity. Lymph node metastases tend to be small and multiple in number. Among patients without clinically apparent lymph node metastasis (staged cN0), pathological examination after lymph node dissection shows occult metastases in up to 40% (1). Accurate staging of the primary tumor and lymph nodes is essential for determination of prognosis and appropriate selection of therapeutic strategies.

Positron emission tomography with ^{18}F-fluor-deoxy-glucose (FDG-PET) provides high contrast resolution and high sensitivity for the detection and staging of a wide variety of malignant diseases. In the head and neck area, some authors have reported a high accuracy of FDG-PET for the detection of primary tumors and lymph node metastases (2-5). However, others have reported a low sensitivity for the detection of lymph node metastases, especially micrometastases (6).
A major drawback of FDG-PET is its relatively limited image quality, expressed in a low spatial resolution and a tendency to produce noisy images. In PET scanning, image quality can be influenced by parameters such as injected dose, acquisition mode and acquisition times. Furthermore, it is reduced by physical limitations, such as annihilation photon non-colinearity, off-axis detector penetration, Compton scatter, and positron range in tissue (7,8). Patient motion may also play a role in adversely affecting image quality.

Image reconstruction
An additional important factor with impact on image quality is the image reconstruction procedure. The choice of an image reconstruction algorithm and its parameters affect contrast resolution, spatial resolution, and level of noise. All three factors influence the measured standardised uptake values (SUVs) and clinical reporting (9). Optimum values depend mainly on scanner characteristics and the level of attenuation. Specifically in the head and neck area, attenuation is much lower as compared to the central body. This results in better count statistics and lower noise levels by default. Less filtering is needed and more iterations may be used without creating unacceptable noise. Standard whole-body reconstruction (SWR) parameters, however, are optimised for the patient as a whole, and disregard particular issues relating to specific body parts such as the head and neck area.

The actual diagnostic benefit of optimised reconstruction algorithms for the head and neck area remains unconfirmed. The aim of this study was to assess the impact of better image quality from optimised head and neck reconstruction (OHR) images on diagnostic yield in the staging of malignancies in the head and neck area.

3.2 Materials and methods

Twenty-eight patients with newly diagnosed malignancies in the oral cavity were included (table 3.1). None of the patients had clinical signs of lymph node metastases, and none had evidence of lymph node metastases on ultrasonographic evaluation of the neck, including fine-needle aspiration of visible lymph nodes when appropriate. None of the patients had a history of diabetes mellitus, and fasting glucose levels were within the normal range.

FDG-PET acquisition

All PET scans were acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens AG, Germany), with a voxel size of 5.15 mm in all directions. One hour after injection of 250 MBq FDG (Tyco-Mallinckrodt, Petten, The Netherlands), emission images were acquired in 3D mode. ^{68}Ge-based transmission imaging was performed for attenuation correction. The acquisition time per bed position was 5 min for emission and 3 min for transmission.

Patients:			**Tumor characteristics at pathology**	
	Sex	Age (y)	Tumor location	Primary
1	F	76	Left oral tongue	T4
2	F	79	Right oral tongue	T2
3	F	55	Right floor of mouth	T1
4	F	55	Ventral floor of mouth	T4
5	F	71	Right floor of mouth	T2
6	F	58	Ventral floor of mouth	T2
7	M	55	Right floor of mouth	T2
8	M	30	Left oral tongue	T2
9	F	66	Left oral tongue	T1
10	F	77	Left oral tongue	T2
11	M	79	Ventral floor of mouth	T2
12	M	47	Ventral floor of mouth	T1
13	F	58	Left oral tongue	T3
14	F	82	Right oral tongue	T2
15	M	53	Right oral tongue	T1
16	M	52	Right oral tongue	T3
17	M	50	Right oral tongue	T1
18	M	64	Ventral floor of mouth	T1
19	F	71	Left oral tongue	T2
20	M	45	Left oral tongue	T1
21	F	59	Right oral tongue	T2
22	M	47	Right oral tongue	T1
23	M	55	Left oral tongue	T3
24	M	66	Left oral tongue (excision biopsy)	T1
25	F	72	Left trigonum retromolare	T2
26	M	79	Right oral tongue	T3
27	M	42	Right oral tongue	T2
28	F	70	Right oral tongue	T3
Avg		61		

Table 3.1
Characteristics of patients included in the study. F = female, M = male.

Image reconstruction

All PET scans were reconstructed using an iterative 2D ordered subset expectation maximisation (OSEM) algorithm (10). For SWR, parameters were left unchanged for normal whole-body imaging (i.e. 2 iterations, 8 subsets and a three-dimensional Gaussian filter of 6 mm). Attenuation correction was based on segmented transmission images. Average reconstruction processing time for a whole-body study was 3 min.

In addition, the head and neck area was reconstructed a second time using the OHR parameters. These parameters were determined beforehand, by three observers who reached a consensus in visual analysis of multiple reconstructions. The best results in respect of spatial resolution, contrast resolution, noise level, blurring and observer confidence were achieved on our system when using 4 iterations, 16 subsets and a three-dimensional Gaussian filter of 5 mm. Attenuation correction was based on non-segmented transmission images. With these settings, the average calculation time for two bed positions was 10 min. Examples of FDG-PET scans of the head and neck area generated with different reconstruction parameters are shown in figure 3.1. In contrast, these values were clearly not suitable for the central body, as illustrated in figure 3.2.

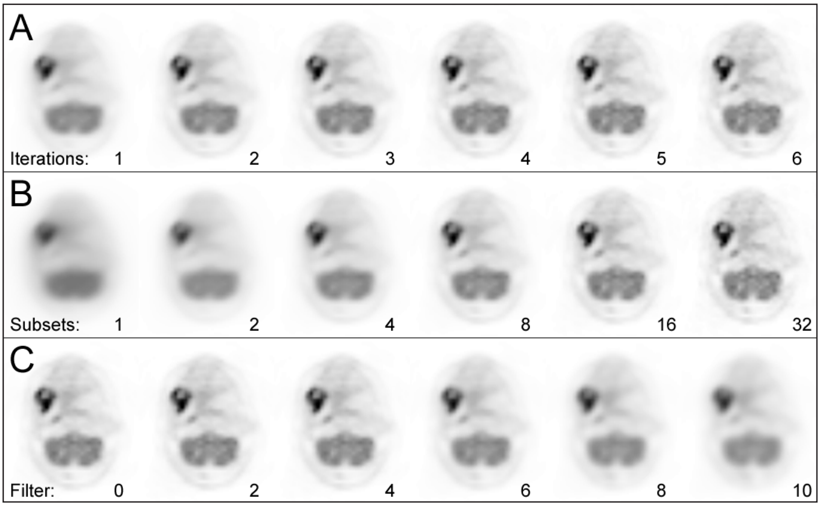


Figure 3.1
Effects of different reconstruction parameters on image quality. From left to right, reconstructions with (A) an increasing number of iterations, using fixed 16 subsets and no filter; (B) an increasing number of subsets, with fixed 4 iterations and no filter; and (C) an increasing Gaussian filter size, with fixed 4 iterations and 16 subsets. Note the changes in visibility of the small lymph node in the dorsal nasopharyngeal wall, as well as better contour detection of the primary tumor.

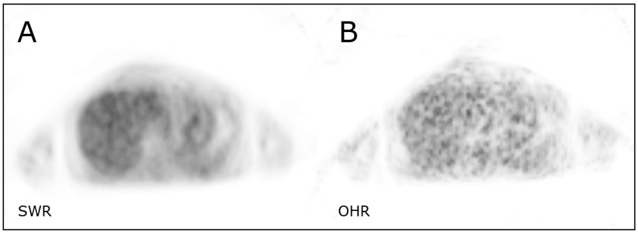


Figure 3.2
The effects of OHR on image quality in more attenuated parts of the body. A transverse slice of a normal upper abdomen is shown. (A) Normal SWR images. (B) OHR images are uninterpretable due to increased noise.

Image analysis

All SWR- and OHR-reconstructed FDG-PET images were read independently by two experienced nuclear medicine physicians blinded to the final histopathological diagnosis. Differences between readers were settled by consensus. The visualisation, location, and additional properties of the primary tumor and the number and location of lesions suspected to represent lymph node metastases were evaluated. SUV_{max} and SUV_{mean} values were measured for all visualised primary tumors, using an automatic 2D region growing algorithm (Hermes, Nuclear Diagnostics, Sweden). Differences in SUV values of the primary tumor were evaluated using a paired *t*-test.

Histopathological verification

After FDG-PET, all patients but one underwent resection of the primary tumor and supraomohyoidal lymph node dissection of the neck within 30 days (average 14 days), providing pathological findings as a gold standard. Patient 26 underwent neck dissection 49 days after FDG-PET. The neck dissection specimens were subjected to standard histopathological examination. All lymph nodes were examined by standard sectioning and haematoxylin and eosin staining. Using the outcome of histopathological staging, the sensitivity, specificity and predictive values for the detection of lymph node metastases by FDG-PET were calculated for both the SWR and the OHR algorithm.

3.3 Results

A summary of the imaging results as compared with histological diagnosis is shown in table 3.2. Subjectively, the OHR images were easier to read. The moderately higher noise level was outweighed by improved spatial resolution and recognition of details in both normal tissue and pathological lesions. Overall, the reviewers judged the OHR images to be of better quality and expressed a higher confidence in the reading.

Primary tumors

In one patient, who had undergone an excision biopsy of the primary tumor, FDG-PET showed no focal pathology. SWR-based images visualised 25 primary tumors in the remaining 27 patients, demonstrating a sensitivity of 92%. Both missed tumors were staged T1 on histopathological examination. On the OHR scans, all 27 primary tumors were visualised, demonstrating a sensitivity of 100%. Figure 3.3.A shows an example of improved visualisation of a primary tumor in the head and neck with OHR images.

As shown in table 3.3, on the SWR images the mean of the SUV_{mean} of all detected primary tumors was 3.8 (range 0.9–8.2) and the mean of the SUV_{max} was 5.2 (range 1.0–11.7). On the OHR images, the mean of the SUV_{mean} was 5.1 (range 1.1–11.1) and the mean of the SUV_{max} was 7.4 (range 1.2–21.3).

Using OHR, the SUV_{mean} increased by 34% ($P<0.0001$) and the SUV_{max} increased by 42% ($P<0.0001$). As shown in figure 3.4, there was a strong linear correlation between the SUVmean values derived from SWR and OHR images (r=0.96), and between the SUVmax values derived from SWR and OHR images (r=0.94).

Table 3.2
The number of primary tumors and pathological lymph nodes visualised with FDG-PET on SWR images and OHR images. The final nodal stage was provided by histopathological examination.

	Primary tumor		Lymph node metastases			
	SWR	OHR	SWR	OHR	Nodal stage	Nodes (diameter)
1	+	+	0	0	pN0	0
2	+	+	0	0	pN0	0
3	+	+	0	0	pN0	0
4	+	+	0	0	pN0	0
5	+	+	0	0	pN0	0
6	+	+	0	0	pN1	1 (2 mm)
7	+	+	0	3	pN0	0
8	+	+	0	1	pN0	0
9	+	+	0	0	pN0	0
10	+	+	0	0	pN0	0
11	+	+	0	0	pN0	0
12	-----	+	0	0	pN1	1 (5 mm)
13	+	+	0	0	pN0	0
14	+	+	0	1	pN1	1 (1 mm)
15	+	+	0	0	pN1	1 (2 mm)
16	+	+	1	1	pN0	0
17	+	+	0	1	pN1	1 (7 mm)
18	+	+	0	0	pN0	0
19	+	+	0	1	pN0	0
20	-----	+	0	0	pN0	0
21	+	+	1	2	pN0	0
22	+	+	0	0	pN0	0
23	+	+	0	0	pN0	0
24	-----	-----	0	0	pN0	0
25	+	+	1	3	pN2b	7 (1 – 8 mm)
26	+	+	0	1	pN2b	5 (3 – 5 mm)
27	+	+	0	0	pN1	1 (4 mm)
28	+	+	0	0	pN1	1 (5 mm)

Lymph nodes

The SWR-based images visualised three lymph nodes in three patients, of which only one was confirmed as a lymph node metastasis on pathological examination. The OHR-based images showed 14 lymph nodes in nine patients (in other words, 11 additional nodes in eight patients). Four of these patients were confirmed to have metastatic lymph nodes in the specified region. Figure 3.3.B illustrates the improved visualisation of lymph node metastases in the head and neck on OHR images.

When calculated on a patient basis (e.g. accuracy for presence of lymph node metastases or not), sensitivity increased from 11% to 44% when using OHR images, while specificity decreased from 89% to 74%. The positive predictive value of the diagnostic procedure increased from 33% to 44% and the negative predictive value increased from 68% to 74%. The diagnostic performance of SWR and OHR images for the detection of lymph node metastases is summarised in table 3.4. With both reconstruction methods, specificity was relatively high, mainly because of a large number of true negatives. All missed lymph node metastases were very small, with a maximum diameter of 5 mm or less.

The values were calculated separately for detection of lymph nodes on an involved neck side basis (e.g. having lymph node metastases on either the left or the right side). When calculated on the basis of the involved neck side, sensitivity increased from 9% to 45% when using OHR images, while specificity decreased from 96% to 87%. The positive predictive value of the diagnostic procedure increased from 33% to 45% and the negative predictive value increased from 81% to 87%.

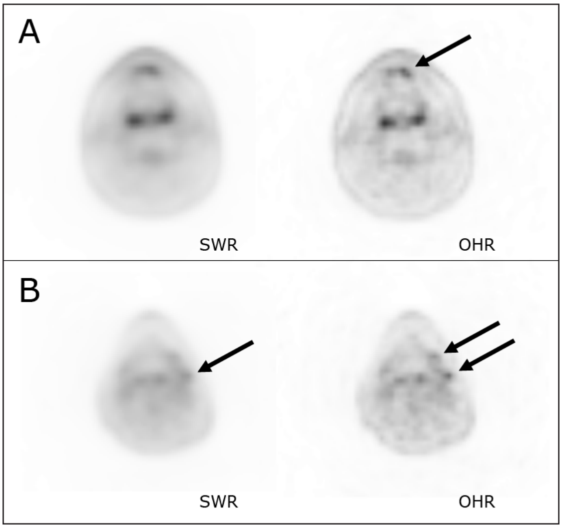


Figure 3.3
Examples illustrating the improvement in image quality between reconstructions with SWR and OHR parameters in the head and neck area. (A) Patient with a T1 carcinoma of the left ventral tongue. On the SWR images the tumor was not distinguishable from physiological tongue muscle activity. It was visible on the OHR images (arrow). (B) Patient with suspicion of one vague lymph node metastasis (arrow) on the SWR images. On the OHR images, multiple nodes were visible (arrows). Note that physiological structures such as the tonsils are also more clearly visualized.

Figure 3.4
The relation between measured SUVs of detected primary tumors in the head and neck area, on SWR and OHR images. There is a strong correlation between the values on SWR and OHR images, though the OHR values are markedly higher than the SWR values. Left, SUV$_{max}$; right, SUV$_{mean}$.

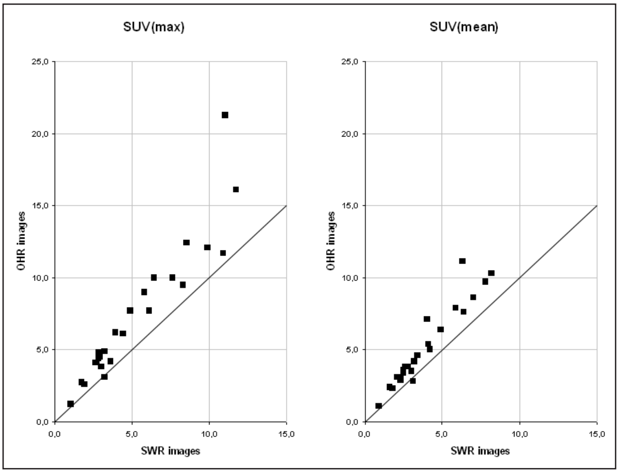


Table 3.3
SUVs measured in visualised primary tumors, in SWR images and in OHR images.

	Primary tumor SUV(max)			**Primary tumor SUV(mean)**		
	SWR	OHR	Increase	SWR	OHR	Increase
1	7,6	10,0	32 %	5,9	7,9	34 %
2	3,9	6,2	59 %	2,8	3,8	36 %
3	2,8	4,4	57 %	2,5	3,6	44 %
4	11,0	21,3	94 %	6,3	11,1	76 %
5	1,0	1,2	20 %	0,9	1,1	22 %
6	1,9	2,6	37 %	1,8	2,3	28 %
7	2,9	4,5	55 %	2,5	3,4	36 %
8	3,2	3,1	-3 %	3,1	2,8	-10 %
9	2,7	4,1	52 %	2,5	3,5	40 %
10	6,4	10,0	56 %	4,0	7,1	78 %
11	11,7	16,1	38 %	8,2	10,3	26 %
12	No SUV measurable					
13	3,2	4,9	53 %	2,6	3,8	46 %
14	6,1	7,7	26 %	4,2	5,0	19 %
15	3,0	3,8	27 %	2,3	2,9	26 %
16	4,4	6,1	39 %	3,2	4,2	31 %
17	1,7	2,7	59 %	1,6	2,4	50 %
18	4,9	7,7	57 %	3,4	4,6	35 %
19	2,8	4,8	71 %	2,3	3,1	35 %
20	No SUV measurable					
21	2,6	4,1	58 %	2,1	3,1	48 %
22	3,6	4,2	17 %	3,0	3,5	17 %
23	9,9	12,1	22 %	7,8	9,7	24 %
24	No SUV measurable					
25	8,3	9,5	14 %	4,9	6,4	31 %
26	10,9	11,7	7 %	7,0	8,6	23 %
27	5,8	9,0	55 %	4,1	5,4	32 %
28	8,5	12,4	46 %	6,4	7,6	19 %
Avg	5,2	7,4	42 %	3,8	5,1	34 %

Overall, when considering primary tumors and lymph node metastases on a patient basis, the FDG-PET diagnosis was improved by using the OHR protocol in six patients (21%). The FDG-PET diagnosis worsened in three patients (11%), in all cases owing to visualisation of lymph nodes without metastasis.

3.4 Discussion

This study shows that a general reconstruction algorithm for whole-body studies may be less suitable for staging malignancies in the head and neck region. Using the SWR protocol, the sensitivity for the presence of lymph node metastases was only 11%, which is considerably lower than the findings reported by some other groups (2–5). The presented results are therefore probably specific for our situation, reflecting, for example, the type of PET scanner and the standard reconstruction parameters used. Factors such as the aforementioned may also partially explain the variable results found in literature.

Using the OHR protocol we observed an increase in sensitivity for the detection of primary tumors, as well as an increase in sensitivity and specificity for the detection of lymph node metastases. From a clinical point of view, the most important improvements when using OHR were the increases in positive predictive value and negative predictive value for the presence of lymph node metastases. Where FDG-PET is used for staging, this will limit the number of missed N+ stages. The benefit is even more clearly expressed by the enhanced performance for the detection of the involved neck side, as this may have a direct impact on patient management when used for selection of lymph node dissection. Although not investigated in the current studies, similar effects may apply for relapse detection, re-staging and therapy monitoring. Perhaps equally as important as the better objective parameters is the possible subjective gain in reviewer confidence in image reading.

These beneficial effects may partially be explained by the higher spatial resolution of OHR-reconstructed FDG-PET images. Current clinical PET scanners have a spatial resolution of 4–7 mm full-width at half-maximum. The detection limit for small lesions is in the range of

Table 3.4
Diagnostic performance of FDG-PET in the detection of lymph node metastases on SWR images and OHR images.

	Patient basis		**Involved neck sides**	
	SWR	ORH	SWR	ORH
Sensitivity (%)	11	44	9	45
Specificity (%)	89	74	96	87
Positive predictive value (%)	33	44	33	45
Negative predictive value (%)	68	74	81	87

4–10 mm, depending on scanner characteristics, localisation of the lesion, and degree of FDG uptake by the lesion (11–13). Although we have not evaluated objective changes in scanner spatial resolution, all lymph node metastases that were additionally detected with OHR had a diameter of 7 mm or smaller. Another important explanation for the benefits is the higher contrast resolution, as a better tumor to background ratio facilitates visual detection of lesions. This effect was quantitatively demonstrated by a statistically significant increase in SUV.

Although the OHR protocol delivers higher sensitivity for the detection of lymph node metastases, it results in a lower specificity due to visualisation of a larger number of small reactive lymph nodes. FDG accumulates in inflammatory cells such as activated granulocytes and macrophages (14), and this inherent characteristic of FDG will be difficult if not impossible to circumvent. Although false positive in oncological staging, from a pathophysiological point of view these lymph nodes are true positive. One can argue that intentional use of suboptimal image quality is not the preferred method to discriminate between reactive lymph nodes and metastases.

Several algorithms can be used for image reconstruction, such as filtered back-projection (FBP) or iterative ordered subset expectation maximisation (OSEM). The choice of image reconstruction algorithm and parameters influences the contrast resolution, spatial resolution and noise level. Nowadays, most investigators acknowledge that iterative reconstruction such as iterative OSEM yields the best image quality (15–18), although some discussion continues (19). We limited this study to optimisation of iterative reconstruction, since FBP reconstruction has nearly been abandoned in clinical practice.
In iterative reconstruction, three parameters are relevant: the number of iterations, the number of iteration subsets and the strength of the filter (20). Essentially, increasing the number of iterations (and/or iteration subsets) improves spatial resolution and contrast resolution at the expense of higher noise levels and increased calculation time. Increasing the filter strength reduces noise at the expense of reduced spatial resolution and contrast resolution. Reconstruction parameters are not generally applicable to all PET systems, as the results depend on camera characteristics, scanning mode, injected dose, the purpose of the scan and the personal preferences of the reviewer. Thus, optimisation of the diagnostic yield of FDG-PET should be pursued by assessing on-site improvement of reconstruction parameters.

Noise in PET images is mainly caused by photon attenuation and scatter, which strongly depend on the amount of attenuation (i.e. the patient’s weight). The standard reconstruction parameters are chosen to render an optimal balance between noise levels, spatial resolution and contrast resolution, for an average patient. In the current study, more noise was observed in the OHR images by the reviewers. The somewhat higher increase in SUV_{max} as compared to SUV$_{mean}$ also reflects increased statistical noise. This did not hinder adequate reporting, as was demonstrated by the improved diagnostic values. An increase in lesion SUV seems to improve the sensitivity for small lesions, especially in the range of the detection limit. This applies not only to FDG scanning: lesions with a poor SUV are likely to be present when using radiopharmaceuticals with a relatively low target-to-background ratio, such as ^{18}F-fluor-deoxy-thymidine (FLT) (21) or ^{18}F-misonidazole. In these cases the improved algorithm may also be of value.

The head and neck region is not the only area where improvement may be achieved. Adaptation of reconstruction parameters may also improve image quality in other situations with relatively low attenuation, e.g. in thin body parts (such as arms, legs and feet) or in patients with low body weight (cachectic patients, young children). An important difference, however, is the variable attenuation within these patients, as compared with the relatively uniform and stable size of the adult head and neck. This implies that a standard reconstruction protocol would be less desirable, but instead different parameters need to be investigated to optimise image quality.

3.5 Conclusion

Routine whole-body PET reconstruction parameters may prove inadequate for the head and neck area. By using image reconstruction parameters adapted to lower photon attenuation, major improvements may be achieved in image quality and diagnostic yield for the detection of small lesions. These improvements were shown to have an impact on staging and patient management. Additional separate reconstruction of the head and neck area with optimised parameters for specific clinical questions is advisable. This may also apply to image reconstruction of other small body parts or to whole-body studies in children.

3.6 References

1. Stuckensen T, Kovacs AF, Adams S, Baum RP: Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J Craniomaxillofac Surg. 2000;28:319-324
2. Bruschini P, Giorgetti A, Bruschini L, Nacci A, Volterrani D, Cosottini M, Ursino F, Mariani G, Fattori B: Positron emission tomography (PET) in the staging of head neck cancer: comparison between PET and CT. Acta Otorhinolaryngol Ital. 2003;23:446-453

3. Teknos TN, Rosenthal EL, Lee D, Taylor R, Marn CS: Positron emission tomography in the evaluation of stage III and IV head and neck cancer. Head Neck. 2001;23:1056-1060

4. Braams JW, Pruim J, Freling NJ, Nikkels PG, Roodenburg JL, Boering G, Vaalburg W, Vermey A: Detection of lymph node metastases of squamous-cell cancer of the head and neck with FDG-PET and MRI. J Nucl Med. 1995;36:211-216

5. Laubenbacher C, Saumweber D, Wagner-Manslau C, Kau RJ, Herz M, Avril N, Ziegler S, Kruschke C, Arnold W, Schwaiger M: Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous-cell carcinomas. J Nucl Med. 1995;36:1747-1757

6. Stoeckli SJ, Steinert H, Pfaltz M, Schmid S: Is there a role for positron emission tomography with 18F-fluorodeoxyglucose in the initial staging of nodal negative oral and oropharyngeal squamous cell carcinoma. Head Neck. 2002;24:345-349

7. Levin CS, Hoffman EJ: Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. Physics in Medicine and Biology. 1999;44:781-799

8. Sanchez-Crespo A, Andreo P, Larsson SA: Positron flight in human tissues and its influence on PET image spatial resolution. Eur J Nucl Med Mol Imaging. 2004;31:44-51

9. Schoder H, Erdi YE, Chao K, Gonen M, Larson SM, Yeung HW: Clinical implications of different image reconstruction parameters for interpretation of whole-body PET studies in cancer patients. J Nucl Med. 2004;45:559-566

10. Hudson HM, Larkin RS: Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imag. 1994;13:601-609

11. DeGrado TR, Turkington TG, Williams JJ, Stearns CW, Hoffman JM, Coleman RE: Performance characteristics of a whole-body PET scanner. J Nucl Med. 1994;35:1398-1406

12. Wienhard K, Dahlbom M, Eriksson L, Michel C, Bruckbauer T, Pietrzyk U, Heiss WD: The ECAT EXACT HR: performance of a new high resolution positron scanner. J Comput Assist Tomogr. 1994;18:110-118

13. Boellaard R, Buijs F, de Jong HW, Lenox M, Gremillion T, Lammertsma AA: Characterization of a single LSO crystal layer high resolution research tomograph. Phys Med Biol. 2003;48:429-448

14. Kubota R, Yamada S, Kubota K, Ishiwata K, Tamahashi N, Ido T: Intratumoral distribution of fluorine-18-fluorodeoxyglucose in vivo: high accumulation in macrophages and granulation tissues studied by microautoradiography. J Nucl Med. 1992;33:1972-1980

15. Hsu CH: A study of lesion contrast recovery for iterative PET image reconstructions versus filtered backprojection using an anthropomorphic thoracic phantom. Comput Med Imaging Graph. 2002;26:119-127

16. Lartizien C, Kinahan PE, Swensson R, Comtat C, Lin M, Villemagne V, Trebossen R: Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. J Nucl Med. 2003;44:276-290

17. Riddell C, Carson RE, Carrasquillo JA, Libutti SK, Danforth DN, Whatley M, Bacharach SL: Noise reduction in oncology FDG PET images by iterative reconstruction: a quantitative assessment. J Nucl Med. 2001;42:1316-1323

18. Etchebehere EC, Macapinlac HA, Gonen M, Humm K, Yeung HW, Akhurst T, Scher HI, Larson SM: Qualitative and quantitative comparison between images obtained with filtered back projection and iterative reconstruction in prostate cancer lesions of (18)F-FDG PET. Q J Nucl Med. 2002;46:122-130

19. Schauwecker DS, Siddiqui AR, Wagner JD, Davidson D, Jung SH, Carlson KA, Hutchins GD: Melanoma patients evaluated by four different positron emission tomography reconstruction techniques. Nucl Med Commun. 2003;24:281-289

20. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA: Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med. 2004;45:1519-1527

21. Buck AK, Halter G, Schirrmeister H, Kotzerke J, Wurziger I, Glatting G, Mattfeldt T, Neumaier B, Reske SN, Hetzel M: Imaging proliferation in lung tumors with PET: 18F-FLT versus 18F-FDG. J Nucl Med. 2003;44:1426-1431

4

FDG-PET in the clinically negative neck in oral squamous cell carcinoma

Bart M. Wensing [1]
Wouter V. Vogel [2]
Henri A.M. Marres [1]
Matthias A.W. Merkx [3]
Ernst J. Postema [2]
Wim J.G. Oyen [2]
Frank J.A. van den Hoogen [1]

Departments of Otorhinolaryngology and head and neck surgery [1], Nuclear medicine [2], and Oral and maxillofacial surgery [3], Radboud University Nijmegen Medical Centre, the Netherlands

Laryngoscope, 2006
Volume 116 (5), pages 809-813

Abstract

With improved diagnostic imaging techniques, it remains difficult to reduce occult metastatic disease in oral squamous cell carcinoma (SCC) to less than 20%. Therefore, supraomohyoid neck dissection (SOHND) still is a valuable staging procedure in these patients.

Methods: Patients with clinically and ultrasonographically staged cN0 SCC of the oral cavity underwent FDG-PET before SOHND. Histologic examination of neck dissection specimens was used as a “gold standard.”

Results: Twenty-eight consecutive patients were included, representing 30 necks. Occult metastatic disease was found in 30% of SOHND specimens. Average diameter of metastatic deposits was 4.3 mm. Sensitivity, specificity, and accuracy of FDG-PET was 33%, 76%, and 63%, respectively.

Conclusions: In patients with cN0 SCC of the oral cavity, FDG-PET does not contribute to the preoperative workup, and cannot replace SOHND as a staging procedure.

4.1 Introduction

The oral cavity is one of the cancer sites in the head and neck that is accompanied by a high incidence of occult regional metastases. Because the presence of cervical lymph node metastasis significantly reduces survival, this important prognostic factor remains an issue for scientific debate. In these discussions, the changing necessity, diagnostic value, and therapeutic role of a selective neck dissection of levels I through III, also called supraomohyoid neck dissection (SOHND), is often addressed. In 1980, Hanley summarizes the opinion at that time about SOHND as a well-recognized, although generally unaccepted form of treatment in patients with oral cavity cancer (1). He compared SOHND with the standard radical neck dissection in 62 patients with oral cavity tumors and found "similar and no worse results for the SOHND." He felt that less than radical neck dissection required further exploration. Since then, the opinion about SOHND has gradually changed (2). SOHND has been popularized over the years and has become a generally accepted staging procedure for the clinically negative neck in patients with oral cancer (3,4). In cases of limited nodal disease on histopathologic examination, some consider it therapeutic (5).

Several studies have shown that even with improved diagnostic imaging techniques, it still remains difficult to reduce occult metastatic disease in oral cavity cancer to less than in 20% of the patients (3,4). The 20% threshold is considered important because elective treatment of the neck remains indicated when the probability of occult neck metastases exceeds 20% (6). Palpation of the neck, with a sensitivity and specificity of 60% to 80%, is not a very accurate way to search for cervical metastases. Computed tomography (CT), magnetic resonance imaging (MRI), ultrasonography (US), sentinel lymph node biopsy (SLNB), and ultrasound-guided fine needle aspiration cytology (USgFNAC) have been shown to be superior to palpation, with the latter being the most accurate way of investigation. However, in patients with a clinically negative neck, even USgFNAC is reported to have a sensitivity of no more than 48% to 76%, most likely explained by the limited tumorload (7,8). USgFNAC of the sentinel node does not seem to increase sensitivity (9). With these results in mind, the risk of occult metastatic disease remains between 15% and 25% (9,10).

Positron emission tomography (PET) with ^{18}F-fluoro-deoxy-glucose (FDG) is increasingly used in preoperative staging of cancer patients. FDG depicts the increased metabolism of malignant cells as compared with normal cells. Reports on the value of FDG-PET in detecting occult metastatic disease have been contradictory, with a reported sensitivity ranging from 0 to 100% and specificity from 92 to 100% (11-13). The added value of FDG-PET to the clinical and ultrasonographic evaluation of the neck in oral carcinoma has not been properly investigated.

We conducted a study to evaluate whether further reduction of occult metastatic disease in oral carcinoma can be achieved by adding FDG-PET scanning to the preoperative workup. With 21% reported occult metastatic disease in oral carcinoma after a negative ultrasonographic investigation in our hospital (10), this could implicate that a diagnostic SOHND would no longer be indicated.

4.2 Methods

This prospective study was performed at the Radboud University Nijmegen Medical Centre between June 2001 and December 2003. All patients planned for SOHND as part of the treatment for a clinically N0 (cN0) squamous cell carcinoma (SCC) of the oral cavity were included in the study, and were offered a FDG-PET scan. A cN0 neck is defined as a neck staged N0 by preoperative palpation by an experienced ear, nose and throat surgeon, oral and maxillofacial surgeon, or radiotherapist and a standard preoperative US performed with USgFNAC on indication. FNAC of enlarged lymph nodes was performed in nodes with a diameter >0.5 cm. In case of a cN-positive neck, the patient was offered a modified radical neck dissection (MRND) and was excluded from the study. The study was approved by the local medical ethics committee. Informed consent was obtained from all patients. No patient experienced cancer within 5 years before this study, or was previously treated by radiotherapy or chemotherapy.

FDG-PET

A dedicated PET-scanner (ECAT-EXACT 47, Siemens/CTI, Knoxville, TN) was used for data acquisition. Before FDG injection, patients fasted for at least 6 hours. Intake of sugar-free liquids was permitted. Immediately before the procedure, the patients were hydrated with 500 mL of water. One hour after intravenous injection of 220 to 250 MBq FDG (Mallinckrodt Medical, Petten, The Netherlands) and 20 mg furosemide, emission and transmission images of the head and neck area were acquired (2–3 bed positions, 5 minutes emission and 3 minutes transmission per bed position). The images were corrected for attenuation and reconstructed using the ordered-subsets expectation maximization (OSEM) algorithm, with 4 iterations, 16 subsets, and a 5 mm 3D Gaussian filter. The reconstructed images were displayed in coronal,

Location	**cT stage**				
	n	T1	T2	T3	T4
Floor of the Mouth	9	3	4	1*	1*
Tongue	17	6	8	3	0
Alveolar proces	2	0	1	0	1

Table 4.1
*Localizations and clinical T stage of the 28 tumors. * = Medial floor of the mouth.*

transverse, and sagittal planes and evaluated by three nuclear medicine physicians. The mean glucose level just before PET imaging was 5.3 mmol/L; three patients had diabetes mellitus (glucose levels 10.3, 6.8, and 3.4 mmol/L). For statistical analysis the sensitivity, specificity and accuracy of FDG-PET were calculated.

Surgical technique and pathologic examination

SOHND was performed as described by Medina et al. (2). In case of enlarged or suspicious nodes found during surgery, frozen stage section (FSS) analysis was performed. If positive, the neck dissection was extended to a MRND. The neck dissection specimen was marked anatomically by the surgeon and sent for pathologic examination consisting of node count, evaluation of malignancy in any of the nodes, and extracapsular spread. Standard sectioning and hematoxylin & eosin staining were performed on all lymph nodes. Lymph node metastases were defined as micrometastasis when a metastatic deposit had a largest diameter of less than 4 mm.

4.3 Results

Thirty patients (15 male, 15 female; median age 60 years, range, 32–84 years) were enrolled. One FDG-PET scan was unevaluable as a result of significant movement of the (diabetic) patient during imaging. One patient had a SOHND 49 days after FDG-PET and was also excluded. Twenty-eight scans could be evaluated (15 women, 13 men), representing 30 SOHNDs. The sites of the primary tumor and T stages are shown in table 4.1. None of the patients had any sign of metastatic spread based on clinical examination and US of the neck. In eight patients, US was supplemented with USgFNAC, which resulted in negative cytology. Additionally, four patients had a CT of the head and neck, all without suspect nodes.

In 26 patients (28 necks), FSS of the neck specimen was performed: 25 (89%) necks were negative and three (11%) were positive. All three necks positive on FSS turned out to be true-positive and one of 25 negative FSS specimens turned false-negative during further pathologic investigation. In the three necks (10% of necks) with a positive FSS of the SOHND specimen, the SOHND was extended to a MRND. No additional metastatic nodes were found in levels IV and V. A total number of 555 nodes was examined in 28 patients. The total number of nodes found in regions I to III varied from 11 to 38. Nine of 30 necks (30%) contained occult metastases in a total of 16 lymph nodes; no bilateral metastases were found. The average diameter of the metastatic deposit was 4.3 mm (range, 0.5–8.0 mm). Five of the malignant nodes showed signs of extracapsular spread, being a prognostic factor of enhanced metastatic spread (3). An overview of the patients and their nodes is presented in table 4.2.

Ultrasound-guided fine needle aspiration cytology

Three of eight patients (38%; three of nine punctured necks: 33%) in whom USgFNAC was negative did have lymph node metastases at pathologic examination of the SOHND specimen; in one case, a node in the punctured level was positive; the other two patients had one and two positive nodes in nonpunctured levels. In patients in whom no USgFNAC was performed, six of 21 SOHND specimens proved to be positive for lymph node metastases (29%).

	Patients			Results			
Nr	Age	M/F	cT-stage	US / FNAC	FDG-PET	PATH	Diameter of metastases (mm)
1	76	F	T4	Neg / np	Neg	Neg	
2	79	F	T2	Neg / np	Neg	Neg	
3	55	F	T1	Neg / np	Neg	Neg	
4a	55	F	T4	Neg / Neg	Neg	Neg	
4b	55	F	T4	Neg / Neg	Neg	Neg	
5	71	F	T2	Neg / np	Neg	Neg	
6	58	F	T2	Neg / np	Neg	1 in I	2 mm
7	55	M	T2	Neg / Neg	2 in II	Neg	
8	30	M	T2	Neg / np	1 in II	Neg	
9	66	F	T1	Neg / np	Neg	Neg	
10	77	F	T2	Neg / np	Neg	Neg	
11	79	M	T2	Neg / np	Neg	Neg	
12	47	M	T1	Neg / Neg	Neg	1 in I	5 mm
13	58	F	T3	Neg / np	Neg	Neg	
14	82	F	T2	Neg / np	1 in II	1 in II	1 mm
15	53	M	T1	Neg / Neg	Neg	1 in I	2 mm
16	52	M	T3	Neg / np	1 in II	Neg	
17	44	F	T2	Neg / Neg	Neg	2 in II	4 and 2 mm
18	50	M	T1	Neg / np	1 in II	1 in II	7 mm
19	64	M	T1	Neg / Neg	Neg	Neg	
20	71	F	T2	Neg / np	1 in II	Neg	
21	45	M	T1	Neg / Neg	Neg	Neg	
22	59	F	T2	Neg / np	2 in II	Neg	
23	47	M	T1	Neg / np	Neg	Neg	
24	55	M	T3	Neg / np	Neg	Neg	
25	66	M	T1	Neg / np	Neg	Neg	
26	72	F	T2	Neg / np	3 in II	6 in II, 1 in I	3x8, 7, 5, 4 and 1 mm
27	42	M	T2	Neg / Neg	Neg	1 in II	4 mm
28a	70	F	T3	Neg / np	Neg	1 in I	5 mm
28b	70	F	T3	Neg / np	Neg	Neg	

Table 4.2
Demographic, clinical, imaging, and pathologic characteristics of the 28 patients. US = ultrasound. FNAC = fine needle aspiration cytology. PATH = pathological examination: number of metastatic glands in level. np = not performed.

FDG-PET
In one patient the FDG-PET scan showed no sign of the primary tumor because this tumor (tongue) was already removed by a radical excision biopsy in a secondary referral center. The other 27 primary tumor sites were correctly depicted. This leads to a sensitivity of 100% for visualizing the primary tumor. In eight cases (29%), FDG-PET showed ipsilateral hot spots suspect for possible metastases; 20 scans representing 22 necks were negative. Patients were on average surgically treated within 14 days (range, 1–30 days; median, 12 days) after receiving their PET scan. Twenty-eight patients underwent 30 neck dissections (two bilateral as a result of a floor of the mouth tumor crossing the midline).

FDG-PET correctly recognized three of nine (33%) afterward pathologically proven positive necks; these positive nodes were found in the correctly depicted level, although one patient had seven positive nodes divided among levels I and II; his PET scan only showed three positive spots in level II. In five patients, FDG-PET showed a suspect node which was not confirmed by pathologic examination; in one of these cases, USgFNAC was performed 3 weeks before FDG-PET. FDG-PET also failed to show any positive nodes in six patients while pathologically present, leading to a sensitivity and specificity of 33% and 76%, respectively, for 30 necks. Accuracy reached 63% for FDG-PET. If SOHND would only be performed when FDG-PET was positive for lymph nodes, the number of SOHNDs would have been reduced by 73% (from 30 to 8). However, this algorithm would result in six of 30 necks with unrecognized (and untreated) occult metastases (20%).

4.4 Discussion

Supraomohyoid neck dissection has shown to be a valid staging procedure in cN0 oral SCC (3). It provides valuable clinical information on the status of the neck at the expense of minimal morbidity (31). CT, MRI, and US lack sensitivity and specificity for diagnosing occult metastatic disease because micrometastases can occur in the absence of morphologic changes in lymph nodes, whereas these changes can be both reactive and metastatic. That is the main reason why USgFNAC is considered to be superior to these imaging modalities, especially when normal-sized and thus nonsuspicious nodes in the first echelons are punctured.

Because the majority of patients in this study did not receive FNAC, this might explain why the 30% of necks with occult metastatic nodes in this study was higher than the initially reported 21% from our institute (10). All metastases in the SOHND specimens were located in the first echelons. Because the metastatic parts of the affected nodes were very small, morphologic changes were unlikely to appear in the majority. Furthermore, CT scanning of the head and neck is no standard preoperative staging technique in patients with a cN0 neck in oral SCC according to the national Dutch guideline on diagnosis and treatment of oral and oropharyngeal cancer.

Several studies have shown that it is possible to reduce occult metastatic disease in oral SCC to approximately 20% (5,10). Because a 20% false-negative rate is considered to be the limit for a wait-and-see policy versus elective treatment of the neck (6), clinicians are expected to be divided into two groups advocating either one of these treatment strategies. However, clinical studies were often retrospective or lacked sufficient power and follow-up time to show a survival benefit for any particular option (33–38). Many surgeons feel that the morbidity of a SOHND is acceptably low and outweighs the risk of being confronted with advanced neck disease after a wait-and-see or wait-and-scan policy (14). Nevertheless, if a diagnostic strategy would further reduce the rate of occult metastasis, SOHND could be abandoned. With this in mind, recent studies have mainly focused on sentinel node biopsy and FDG-PET scanning of the neck.

Sentinel lymph node biopsy (SLNB) in cutaneous malignant melanoma has been demonstrated to be a minimally invasive technique with a high degree of accuracy in detecting occult metastatic disease. Morton et al. (40). reported a false-negative rate of SLNB in cN0 cutaneous malignant melanoma of the head and neck area of less than 1%. SLNB might possibly identify those patients who harbor occult metastatic disease in cN0 oral SCC. Ross et al. describe the preliminary results of a multicenter trial in patients with T1 to 2N0 oral cavity/oropharynx carcinoma (15). In 93% of 134 cases, a sentinel node could be identified. In 55 cases, SLNB was combined with an elective neck dissection. Occult metastatic disease was present in 34%. Sensitivity of this technique was reported to be 93% in all oral SCC. In floor of the mouth SCC only, however, identification of the SLN was possible in 86% (n = 43) and sensitivity reached 80%. Besides selection of patients for (additional) neck surgery, histopathologic evaluation of the sentinel node might limit cost and time-consuming pathologic evaluation by performing step serial sectioning and immunohistochemistry of the sentinel nodes that are most suspected to harbor metastases. However, skip metastases might prove to be a problem (16). Difficulties identifying level I sentinel lymph nodes in oral SCC have been reported (15). In our series, 31% of the metastatic nodes were located in level I. Thus, questions about whether SLNB offers a significant advantage over SOHND still remain. Unresolved issues regarding this technique are the feasibility of and efficacy in multiple-level sentinel nodes and the cost-effectiveness.

Furthermore, underestimation of occult metastatic disease with standard sectioning and hematoxylin & eosin staining is likely to be present in approximately 5% to 8% (17,18). It is unclear if these micrometastases need more than a SOHND or if SOHND can be considered adequate treatment in these cases. Potential overtreatment of such limited disease with a MRND needs to be subject of clinical studies as stated by Pitman et al. in a review on SLNB in head and neck cancer (19). Nieuwenhuis et al. investigated the role of USgFNAC of the sentinel node in cT1 to 2N0 oral/oropharyngeal carcinoma (9). Although the sentinel node could be identified and aspirated in 38 of 39 patients, it did not decrease the false-negative rate compared with USgFNAC alone. The additional value of sentinel node cytology was thus questioned.

Considerable attention has been paid to FDG-PET in the evaluation of the N0 neck in head and neck cancer. In our study, we have included only patients without evidence of metastatic neck disease in preoperative evaluation consisting of at least palpation and US and who were thus scheduled for surgical treatment, including a SOHND. Sensitivity in this selected patient group is insufficient to refrain from SOHND on the basis of FDG-PET.

Initial reports about FDG-PET for lymph node staging of the neck in head and neck SCC were very optimistic with sensitivity and specificity values between 72% to 91% and 88% to 98%, respectively (26,47). Stuckensen et al. performed a prospective study in 106 patients with oral SCC and compared FDG-PET, US, CT, and MRI with histologic results of the neck dissection specimen (24). Only 48 of 106 patients included in the study were staged cN0. FDG-PET had a sensitivity of 70% and a specificity of 82%. There was a correlation with the size of the lymph node metastases. Most of the metastases larger than 12 mm were detected, but less than half of those smaller than 6 mm were detected. Therefore, it is likely that sensitivity and specificity in the cN0 group were worse than those reported for the complete group.

	No. of patients				**Diagnostic value of PET**		
Study	Total in study	Palpably N0	N0 palpably / radiographically	Pathology	Sens.	Spec.	Accuracy
Myers et al.	14	14	8 / 12	*	78%	100%	92%
Civantos et al.	18	18	18 / 18	IHSN	30%	100%	61%
Hyde et al.	19	18	18 / *20	IHSN	0%	100%	78%
Brouwer et al.	15	15	12 / 12	HE	67%	92%	87%
Wensing et al.	28	28	28 / 30	HE	33%	76%	63%
Total	94	93	84 / 92	---	38%	92%	74%
Corrected data for 84 patients palpably and radiographically N0					† 25%	91%	72%

Table 4.3
*Comparison of studies considering the use of FDG-PET to find metastatic spread in patients with head and neck carcinoma (laryngeal and hypopharyngeal cancer excluded) regarding the clinically N0 neck. * = Could not be retrieved from study. † = Corrected percentages are based on number of treated necks (108 of Braams et al. because no sensitivity could be estimated in the two remaining patients). IHSN = immunohistochemy of sentinel node only. HE = hemotoxylin and eosin staining.*

In FDG-PET studies specifically addressing patients with oral cavity or oropharyngeal cancer with a cN0 neck, huge variations in sensitivity (from 0–100%) and specificity (92–100%) are reported (11–13,20). Brouwer et al. showed that in the studies using routine histopathologic workup, a much higher sensitivity (67–100%) for the detection of occult lymph node metastases was found in comparison to the studies in which step sectioning and immunohistochemistry as part of the sentinel node procedure was performed (11,22). In the latter studies, a sensitivity of only 0% to 40% was found (20). They point out that the histopathologic method used seems to be the most important factor for the differences in sensitivity. However, we feel that patient selection and differences in inclusion criteria are other major reasons for these differences. In the "less sensitive studies", only patients were included who were clinically as well as ultrasonographically N0, and in the "sensitive studies", patients were included who had a clinically negative neck but often had radiologic evidence of metastatic disease (12–14,20). If the results of the studies in table 4.3 are combined, while excluding patients with radiologic evidence of metastases, overall sensitivity drops from 38% to 25% (84 patients, 92 necks). This is in line of the diagnostic yield of FDG-PET in our study of patients with lymph node metastases all smaller than 8 mm.

4.5 Conclusions

FDG-PET does not lower the false-negative rate of occult lymph node metastases in patients with an oral SCC and a clinically and US-negative neck below the clinically required 20% of patients. Therefore, FDG-PET has no added value in the preoperative workup. FDG-PET alone or in combination with US (±FNAC) cannot replace SOHND as a staging procedure.

4.6 References

1. Snow GB, Annyas AA, van Slooten EA, Bartelink H, Hart AA. Prognostic factors of neck node metastasis. Clin Otolaryngol 1982;7(3):185-92.
2. Leemans CR, Tiwari R, Nauta JJ, van der Waal I, Snow GB. Recurrence at the primary site in head and neck cancer and the significance of neck lymph node metastases as a prognostic factor. Cancer 1994;73(1):187-90.
3. Hanley DJ. Supraomohyoid neck dissection.Br J Plast Surg. 1980;33(1):136-7.
4. Spiro JD, Spiro RH, Shah JP, Sessions RB, Strong EW. Critical assessment of supraomohyoid neck dissection. Am J Surg 1988;156(4):286-9.

5. Medina JE, Byers RM. Supraomohyoid neck dissection: rationale, indications, and surgical technique. Head and Neck 1989;11(2):111-22.

6. Majoufre C, Faucher A, Laroche C et al. Supraomohyoid neck dissection in cancer of the oral cavity. Am J Surg 1999;178(1):73-7.

7. Brazilian Head and Neck Cancer Study Group.Results of a prospective trial on elective modified radical classical versus supraomohyoid neck dissection in the management of oral squamous carcinoma. Am J Surg 1998;176(5):422-7.

8. Shah JP. Patterns of cervical lymph node metastasis from squamous carcinomas of the upper aerodigestive tract. Am J Surg 1990;160(4):405-9.

9. Manni JJ, van den Hoogen FJ. Supraomohyoid neck dissection with frozen section biopsy as a staging procedure in the clinically node-negative neck in carcinoma of the oral cavity. Am J Surg 1991;162(4):373-6.

10. Byers RM. Modified neck dissection. A study of 967 cases from 1970 to 1980. Am J Surg 1985;150(4):414-21.

11. Kerrebijn JD, Freeman JL, Irish JC et al. Supraomohyoid neck dissection. Is it diagnostic or therapeutic? Head and Neck 1999;21(1):39-42.

12. Kowalski LP, Carvalho AL. Feasibility of supraomohyoid neck dissection in N1 and N2a oral cancer patients. Head and Neck 2002;24(10):921-4.

13. Weiss MH, Harrison LB, Isaacs RS. Use of decision analysis in planning a management strategy for the stage N0 neck. Arch Otolaryngol Head Neck Surg 1994;120(7):699-702.

14. Baatenburg de Jong RJ, Rongen RJ, Lameris JS, Harthoorn M, Verwoerd CD, Knegt P. Metastatic neck disease. Palpation vs ultrasound examination. Arch Otolaryngol Head Neck Surgery 1989;115(6):689-90.

15. Feinmesser R, Freeman JL, Noyek AM, Birt BD. Metastatic neck disease. A clinical/radiographic/pathologic correlative study. Arch Otolaryngol Head Neck Surg 1987;113(12):1307-10.

16. Van den Brekel MW, Castelijns JA, Croll GA et al. Magnetic resonance imaging vs palpation of cervical lymph node metastasis. Arch Otolaryngol Head Neck Surg 1991;117(6):663-73.

17. Knappe M, Louw M, Gregor RT. Ultrasonography-guided fine-needle aspiration for the assessment of cervical metastases. Arch Otolaryngol Head Neck Surg 2000;126(9):1091-6.

18. van den Brekel MW, Castelijns JA, Stel HV et al. Occult metastatic neck disease: detection with US and US-guided fine-needle aspiration cytology. Radiology 1992;180(2):457-61.

19. Takes RP, Righi PD, Meeuwis CA et al. The value of ultrasound with ultrasound-guided fine-needle aspiration biopsy compared to computed tomography in the detection of regional metastases in the clinically negative neck. Int J Radiat Oncol Biol Phys 1998;40(5):1027-32.

20. Righi PD, Kopecky KK, Caldemeyer KS, Ball VA, Weisberger EC, Radpour S. Comparison of ultrasound-fine needle aspiration and computed tomography in patients undergoing elective neck dissection. Head and Neck 1997;19(7):604-10.

21. Nieuwenhuis EJ, Castelijns JA, Pijpers R et al. Wait-and-see policy for the N0 neck in early-stage oral and oropharyngeal squamous cell carcinoma using ultrasonography-guided cytology: is there a role for identification of the sentinel node? Head and Neck 2002;24(3):282-9.

22. Van den Brekel MW, Castelijns JA, Reitsma LC, Leemans CR, van der Waal I, Snow GB. Outcome of observing the N0 neck using ultrasonographic-guided cytology for follow-up. Arch Otolaryngol Head Neck Surg 1999;125(2):153-6.

23. van den Hoogen FJA, Westerbeek BC, Joosten FB, Bruaset I, Marres HA. Supraomohyoid neck dissection in the ultrasonographically and clinically negative neck in oral carcinoma. Clin Otolaryngol 2001;26:340.

24. van Wilgen CP, Dijkstra PU, van der Laan BF, Plukker JT, Roodenburg JL. Shoulder complaints after nerve sparing neck dissections. Int J Oral Maxillofac Surg 2004;33(3):253-7.

25. Spiro RH, Strong EW. Epidermoid carcinoma of the oral cavity and oropharynx: elective vs therapeutic radical neck dissection as treatment. Arch Surg 1973;107(3):382-4.

26. Nahum AM, Bone RC, Davidson TM. The case for elective prophylactic neck dissection. Laryngoscope 1977;87:588-99.

27. Cunningham MJ, Johnson JT, Myers EN, Schramm VL Jr., Thearle PB. Cervical lymph node metastasis after local excision of early squamous cell carcinoma of the oral cavity. Am J Surg 1986;152(4):361-6.

28. VandenBrouck C, Sancho-Garnier H, Chassagne D, Saravane D, Cachin Y, Micheau C. Elective versus therapeutic radical neck dissection in epidermoid carcinoma of the oral cavity: results of a randomized clinical trial. Cancer 1980;46(2):386-90.

29. Kligerman J, Lima RA, Soares JR et al. Supraomohyoid neck dissection in the treatment of T1/T2 squamous cell carcinoma of oral cavity. Am J Surg 1994;168(5):391-4.

30. Fakih AR, Rao RS, Patel AR. Prophylactic neck dissection in squamous cell carcinoma of oral tongue: a prospective randomized study. Semin Surg Oncol 1989;5(5):327-30.

31. Andersen PE, Cambronero E, Shaha AR, Shah JP. The extent of neck disease after regional failure during observation of the N0 neck. Am J Surg 1996;172(6):689-91.

32. Morton DL, Wen DR, Foshag LJ, Essner R, Cochran A. Intraoperative lymphatic mapping and selective cervical lymphadenectomy for early-stage melanomas of the head and neck. J Clin Oncol 1993;11(9):1751-6.

33. Ross GL, Soutar DS, MacDonald GD et al. Sentinel node biopsy in head and neck cancer: preliminary results of a multicenter trial. Ann Surg Oncol 2004;11(7):690-6.

34. Ferlito A, Shaha AR, Rinaldo A, Pellitteri PK, Mondin V, Byers RM. “Skip metastases” from head and neck cancers. Acta Otolaryngol 2002;22(7):788-91.

35. Kitagawa Y, Ohgami M, Fujii H et al. Laparoscopic detection of sentinel lymph nodes in gastrointestinal cancer: a novel and minimally invasive approach. Ann Surg Oncol 2001;8(9 supp.):86-9.

36. Barranger E, Cortez A, Commo F et al. Histopathological validation of the sentinel node concept in cervical cancer. Ann Oncol 2004;15(6):870-4.

37. Ambrosch P, Brinck U. Detection of nodal micrometastases in head and neck cancer by serial sectioning and immunostaining. Oncology (Huntingt) 1996;10(8):1226-9.

38. Enepekides DJ, Sultanem K, Nguyen C, Shenouda G, Black MJ, Rochon L. Occult cervical metastases: immunoperoxidase analysis of the pathologically negative neck. Otolaryngol Head Neck Surg 1999;120(5):713-7.

39. Woolgar JA. Micrometastasis in oral/oropharyngeal squamous cell carcinoma: incidence, histopathological features and clinical implications. Br J Oral Maxillofacial Surg 1999;37(3):181-6.

40. Pitman KT, Ferlito A, Devaney KO, Shaha AR, Rinaldo A. Sentinel lymph node biopsy in head and neck cancer. Oral Oncol 2003;39(4):343-9.

41. Kau RJ, Alexiou C, Laubenbacher C, Werner M, Schwaiger M, Arnold W. Lymph node detection of head and neck squamous cell carcinomas by positron emission tomography with fluorodeoxyglucose F 18 in a routine clinical setting. Arch Otolaryngol Head Neck Surg 1999;125(12):1322-28.

42. Stuckensen T, Kovacs AF, Adams S, Baum RP. Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J Craniomaxillofac Surg 2000;28(6):319-24.

43. Brouwer J, de Bree R, Comans EF, Castelijns JA, Hoekstra OS, Leemans CR. Positron emission tomography using [(18)F]fluorodeoxyglucose (FDG-PET) in the clinically negative neck: is it likely to be superior? Eur Arch Otorhinolaryngol 2004;261(8):479-83.

44. Braams JW, Pruim J, Freling NJ et al. Detection of lymph node metastases of squamous-cell cancer of the head and neck with FDG-PET and MRI. J Nucl Med 1995;36(2):211-6.

45. Myers LL, Wax MK, Nabi H, Simpson GT, Lamonica D. Positron emission tomography in the evaluation of the N0 neck. Laryngoscope 1998;108(2):232-36.

46. Myers LL, Wax MK. Positron emission tomography in the evaluation of the negative neck in patients with oral cavity cancer. J Otolaryngol 1998;27:342-7.

47. Hyde NC, Prvulovich E, Newman L, Waddington WA, Visvikis D, Ell P. A new approach to pre-treatment assessment of the N0 neck in oral squamous cell carcinoma: the role of sentinel node biopsy and positron emission tomography. Oral Oncol 2003;39(4):350-60.

48. Stoeckli SJ, Steinert H, Pfaltz M, Schmid S. Is there a role for positron emission tomography with 18F-fluorodeoxyglucose in the initial staging of nodal negative oral and oropharyngeal squamous cell carcinoma. Head and Neck 2002;24(4):345-49.

49. Civantos FJ, Gomez C, Duque C et al. Sentinel node biopsy in oral cavity cancer: correlation with PET scan and immunohistochemistry. Head and Neck 2003;25(1):1-9.

5

Correction of an image size difference between PET and CT improves image fusion of dedicated PET and CT

Wouter V. Vogel [1]
Jorn A. van Dalen [1]
Dominic A.X. Schinagl [2]
Johannes H.A.M. Kaanders [2]
HenkJan Huisman [3]
Frans H.M. Corstens [1]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1], Radiation oncology [2], and Radiology [3],
Radboud University Nijmegen Medical Centre, the Netherlands

Nuclear Medicine Communications, 2006
Volume 27 (6), pages 515-519.

Abstract

Clinical work in software PET/CT image fusion in our institute raised suspicion that the image sizes of PET and CT differred slightly from each other, thus rendering the images suboptimal for image fusion. The aim of this study was to evaluate the extent of the relative image size difference between PET and CT and the impact of the correction of this difference on the accuracy of image fusion.

Methods: The difference in real image size between PET and CT was evaluated using a phantom study. Subsequently, 13 patients with cancer in the head/neck area underwent both CT and FDG-PET in a custom-made mask for external beam radiotherapy, with multimodality markers for positional reference. The image size of PET relative to CT was determined by evaluating the distances between the markers in multiple directions in both scans. Rigid-body image fusion was performed using the markers as landmarks, with and without correction of the calculated image size difference.

Results: Phantom studies confirmed a difference in real image size between PET and CT, caused by an absolute error in PET image size calibration. The clinical scans demonstrated an average relative difference in image size of 2.0% in the transverse plane and 0.8% along the longitudinal axis, the PET images being significantly smaller. Image fusion using original images demonstrated an average registration error of 2.7 mm. This error was decreased to 1.4 mm after size correction of the PET images, a significant improvement of 48% ($P < 0.001$).

Conclusions: A significant deviation in PET image size may occur, either as a real image size deviation or as a relative difference from CT. Although possibly not clinically relevant in normal diagnostic procedures, correction of such a difference benefits image fusion accuracy. Therefore, it is advisable to calibrate the PET image size relative to CT before performing high-accuracy rigid-body image fusion.

5.1 Introduction

Image fusion of positron emission tomography (PET) and computed tomography (CT) can improve the diagnostic value and accuracy in oncological imaging of the head and neck area (1-3). Image fusion may also be applied to incorporate functional information in external beam radiation treatment (4,5). When performing image fusion, a high accuracy in anatomical registration of the images is required, because incorrect registration may induce diagnostic errors, such as erroneous localization or characterization of the lesions (6). In particular, when using image fusion for the definition of target volumes in intensity-modulated radiation therapy (IMRT), the required accuracy is high as the error in dose delivery is in the range of only 2-3 mm (7). Errors in image registration may influence the outcome of therapy and the level of complications of external beam radiation therapy.

For software image fusion of dedicated PET and CT, an accuracy of better than 2 mm has been demonstrated using phantoms (8). The accuracy that can be achieved in patients will probably be lower as a result of complicating factors, such as small positioning errors, motion artefacts, the time interval between scans and limited comparability between scans due to visualization of different structures and processes on PET and CT. Furthermore, differences may exist in image size. In this article, real image size is defined as the discrepancy between the measured size of an object on an image and the true size of that object. Furthermore, relative differences in image size may occur between scanning modalities. The image size is often not considered in rigid-body image fusion, as – in general – the scanner image sizes are fixed and validated for both PET and CT.

During software PET/CT image fusion with multimodality markers for IMRT planning of the head and neck area, we observed a small systematic scaling difference between CT and PET images. Patients were slightly smaller on PET images than on CT images. An example is shown in figure 5.1.A. It was suspected that the image size of the CT scanner and/or PET scanner was inaccurate. Attempts were made to correct these errors with patient-specific manual or automatic scaling procedures, but the results were variable and were complicated by other factors, such as small positioning differences and deformations.

It was hypothesized that correction with an objectively determined systematic scaling factor would provide a better and more elegant solution. The purposes of this study were to determine the extent of image size differences between PET and CT, to evaluate the impact of this problem on image fusion accuracy and, if needed, to determine a systematic correction factor.

5.2 Materials and methods

Phantom experiment

A linear phantom, 50 cm in length, with 11 multimodality markers positioned at 5 cm intervals, was used to estimate the real image sizes of PET and CT, and to detect the linearity of deviations. Each marker consisted of two glass capillaries positioned under a 90º angle, filled with either iodine-containing X-ray contrast solution or diluted ^{18}F-fluor-deoxy-glucose (FDG) solution. The visualization of a marker is demonstrated in figure 5.2. The phantom was scanned in three directions (i.e. x, y and z) in both PET and CT. The distances between the centres of all marker pairs were measured in three directions in both imaging modalities. Differences in the distances between corresponding marker pairs on PET and CT were evaluated throughout the field of view of the scanners to determine the linearity of deviations. The inter-operator variation of the manual localization of the centre of the markers was evaluated by the analysis of 12 markers by two operators in a separate session.

Clinical experiment

Thirteen patients with newly diagnosed malignancy in the head and neck area were included. All patients had squamous cell carcinoma of the oral cavity or larynx, and were candidates for external beam therapy. None of the patients had a history of diabetes mellitus, and fasting glucose levels were within the normal range. In addition to a standard planning CT, an ^{18}F-

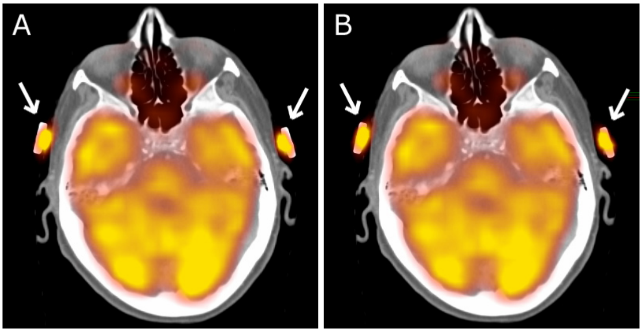


Figure 5.1
Software image fusion of PET and CT. Shown is a slice through the head, at the level of two multimodality fiducial markers positioned in front of the ears. (A) Original images: The markers on PET are closer to each other than on CT, hence the PET image is somewhat smaller than the CT image. (B) After correction for image size differences: The markers are centered correctly.

FDG PET scan was performed to provide biological parameters for optimized target volume definition. Both CT and ^{18}F-FDG PET were acquired in a personalized custom-moulded mask to provide identical positioning. Four multimodality markers were attached to each mask at corner positions outside the target field (figure 5.2).

Image acquisition

CT scans were acquired using a multislice spiral CT scanner (Marconi AcQsim, Marconi Corporation, Cleveland, Ohio, USA). The scanning parameters comprised a scan range from the skull base to the lung top, intravenous contrast in the arterial phase, 100 mAs, 130 kV. The pixel size of the images was 0.938 mm^2 in the transaxial plane. The voxel size in the axial direction was defined by a slice thickness of 3 mm.

PET scans were acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). An activity of 250 MBq of ^{18}F-FDG was injected intravenously. The scans were acquired 1 h post-injection, using three-dimensional emission for 6 min per bed position, and employing attenuation correction based on two-dimensional germanium-68 transmission images for 2 min per bed position. All PET scans were reconstructed using an iterative two-dimensional ordered subset expectation maximization (OSEM) algorithm (9) using four iterations, 16 subsets and a three-dimensional Gaussian filter of 5 mm. A zoom factor of 1.5 was applied to generate voxels with a size of 3.432 mm in all directions.

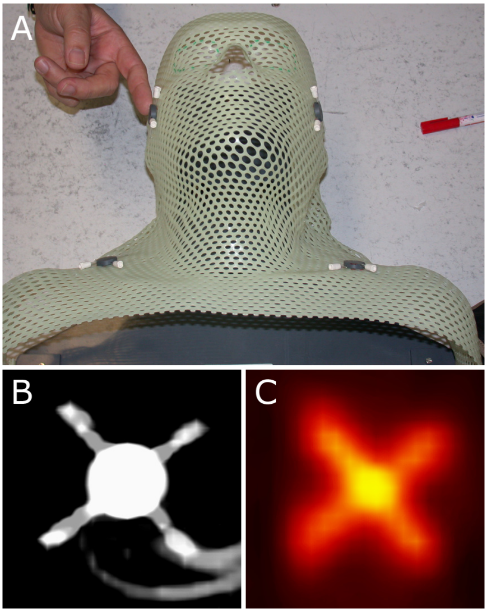


Figure 5.2
Example of the multi-modality markers used for landmark registration and image fusion accuracy evaluation. (A) Markers as placed on the mask. (B) Marker as seen on CT filled with iodine contrast. (C) Marker as seen on PET filled with ^{18}F-FDG solution. The center of the markers could be determined with an inter-operator variability well below 1 mm on both CT and PET images.

Evaluation of image size difference

The difference in real image size between PET and CT was evaluated by measuring the distances between the centres of multiple markers in the transverse and axial planes for both imaging modalities. For each patient, four markers were placed in a rectangular configuration, providing 26 marker pairs in the transverse plane and 26 pairs in the axial direction. Three markers in three separate patients were excluded from evaluation because of poor visibility. Thus, 24 evaluable marker pairs in the transverse plane and 23 pairs in the axial direction were used for the analysis. The relative difference in pixel size between PET and CT was calculated using the formula: $(\text{Distance}_{\text{PET}} - \text{Distance}_{\text{CT}}) / \text{Distance}_{\text{CT}} * 100\%$.

Impact on image fusion

Image fusion was performed twice, with the original PET images and with PET images that had been corrected for the image size difference relative to CT by adaptation of the pixel size in the DICOM header of the source files. Rigid-body landmark-based registration of PET images to CT images was performed with in-house-developed software, based on the visualization toolkit VTK (10). First, the centre of all markers was identified manually. Subsequently, rigid-body registration, i.e. based on three translation and three rotation parameters, of PET to CT images was performed automatically by minimizing the sum of the square distances between corresponding marker centres (11). Scaling of the images in the image registration software was not necessary, because image size corrections had already been performed in the DICOM source files when appropriate.

The accuracy of the image registration procedure was evaluated mathematically by determining the remaining positional difference between corresponding markers on PET and CT after image registration. Statistical analysis of the difference in image registration accuracy between the two image sets was performed using a two-sided paired t-test. The level of significance was set at 0.05.

5.3 Results

The inter-operator variability in the determination of the centre of the markers was well below 1 mm in both types of scans. For CT images, the average difference between the two operators was 0.54 mm (S.D. 0.28 mm). For PET images, the average difference was 0.42 mm (S.D. 0.19 mm).

Image size difference

In the phantom experiment, in the horizontal direction (x-axis), the measured distance between the outermost markers on PET was 490.4 mm. On CT images, the measured distance was 499.2 mm, a relative difference from PET of 1.8%. In the vertical direction (y-axis), the respective distances were 499.0 and 490.1 mm, a relative difference of 1.8%. In the longitudinal direction, the distances were 500.5 and 497.2 mm, a relative difference of 0.7%. For both PET and CT, the markers were distributed evenly along the phantom in all directions. The difference in distance between marker pairs on PET and CT was detected throughout the field of view, indicating a linear image size difference.

In the series of clinical scans, in the transverse plane, a significant average relative difference in distance between markers on PET and CT of 2.0% was observed (range 0.2 – 3.7%, S.D. 0.76%), the patient being smaller on the PET images than on the CT images in all cases. In the axial direction, an average difference of 0.8% (range -0.6 to 2.5%, S.D. 0.65%) was found, the patient being smaller on the PET images than on the CT images in the majority of cases.

To compensate for the detected differences in the clinical series, the pixel size of the PET images was increased by 2.0% to 3.501 mm in the transverse plane, and increased by 0.8% to 3.459 mm along the longitudinal axis, by adaptation of pixel size values in the DICOM file header. Repeated evaluation using the corrected PET images showed an image size difference of 0.2% (range -1.4 – 1.8%, S.D. 0.69%) in the transverse plane, and a difference of 0.0% (range -1.5 – 1.7%, S.D. 0.68%) in the axial direction. Thus, a significant relative difference in image size between PET and CT could no longer be demonstrated.

Accuracy of image fusion

Image fusion using the original uncorrected PET images demonstrated an average mathematical registration error of 2.7 mm (range 0.8 – 5.5 mm) at the position of the markers. When image fusion was performed using PET images corrected for relative image size differences, the error was 1.4 mm (range 0.3 – 3.8 mm). This represented a significant decrease in the error in image fusion at the location of the markers of 48% ($P < 0.001$). Figure 5.1.B shows a fused image after pixel size correction.

5.4 Discussion

The suspected relative difference in real image size between PET and CT was confirmed using both phantom studies and clinical scans. In the clinical scans, a series of separate measurements over multiple patients demonstrated the deviation with statistical significance. The manual identification of the centre of the markers was excluded as a possible source of error in the analysis procedure, because the inter-operator variability was negligible.

The results of the phantom measurements and the clinical scans were concordant (the size differences in the phantom measurements were well within the standard deviations of the clinical series). The deviation was shown to be linear, and thus the results from the phantom study and the clinical data were theoretically exchangeable. The image size difference derived from the clinical series was considered to be the most accurate, because statistical analysis of a series of measurements is less prone to bias in the manual evaluation of markers than a single measurement in the phantom experiment. Therefore, the average image size difference derived from the clinical series was applied as the correction factor. This approach was validated by repeated measurements after correction, which no longer demonstrated significant image size differences.

The observed difference in image size between PET and CT was larger in the transverse plane than in the axial direction (2.0% versus 0.8%). The exact cause of this discrepancy remains unclear. There is no apparent reason why image size calibration of PET images would be more difficult in one direction than in another.

The relative difference in real image size between PET and CT was caused by an absolute error in calibration of the PET image size. With a pixel size and effective resolution (full-width at half-maximum) in the range of 5 mm in PET imaging, accurate image size calibration may be difficult. For example, to detect a deviation of 1%, the difference to be found in a marker distance over a length of 20 cm is only 2 mm, which is well below the resolution. Therefore, the procedure to determine the image size of a PET scanner will always be less accurate, relative to a CT scanner. Furthermore, in IMRT planning, CT is the gold standard by default because planning of the radiation fields depends on the electron density information derived from CT. For these reasons, we have adapted the PET images to match CT, regardless of the possibility of a hypothetical small remaining error in CT real image size.

Correction of the image size differences alone will not result in perfect image fusion. Other causes of inaccuracies remain, such as slight deformation of the mask between scans (as a result of small patient positioning differences) and patient motion during scanning (for example swallowing). These factors, in combination with the sampling errors of the marker locations, contribute to the detected remaining error in rigid-body landmark-based image registration.

It can be argued that image size correction should be applied to all PET scanning, but the relevance of an error of this small magnitude in normal diagnostic imaging is probably negligible. Image size corrections need to be advocated only when a very high accuracy is required, such as in image fusion for IMRT planning in the head and neck area.

Several approaches to rigid-body image registration are available. Examples include manual procedures, automatic methods based on mutual information or the iterative closest point algorithm, and landmark-based registration. We used the latter method for our accuracy evaluation, as landmark registration is a robust and accurate technique when reliable landmarks are available (11). It seems obvious that other rigid-body registration techniques will benefit similarly from the correction of image size differences. Theoretically, when non-rigid transformations are used, there is no need for additional image size corrections.

Multiple approaches may be available to correct the standard image size of PET images. Some PET scanners may allow easy adaptation of the image size on the machine itself, most likely as a parameter in the reconstruction algorithm. Other options include adaptation of the DICOM file that is transported to the fusion software, or adaptation in the fusion software itself. Not all available approaches may support separate adaptation for the transverse and axial directions, as was needed in our specific situation. Otherwise, there are no rational arguments to prefer one approach over another.

As an alternative solution to systematic correction of differences in image size, patient-specific scaling may be advocated as it is relatively easy to perform. Image fusion software generally supports scaling, either manual or automatic. Scaling is even considered as a routine procedure in image fusion with magnetic resonance imaging (MRI), which suffers from spatial distortions and size deviations because of magnetic field inhomogeneities. However, when spatial inhomogeneities are absent, we consider this approach to be suboptimal, as the applied size corrections will be influenced by incidental regional variations, such as positioning differences. The application of *ad hoc* scaling factors seems a less rational procedure, when the exact correction values can be derived by relatively simple measurements as presented in this paper.

It seems unlikely that the correction parameters presented in this article can be transferred to other PET systems in general. Variations may occur between systems, especially when using scanners from different manufacturers. Therefore, the extent of image size differences needs to be assessed locally. Hybrid PET/CT systems may also suffer from relative image size differences between PET and CT, as all currently available hybrid systems consist of dedicated PET and CT scanners placed in-line. Therefore, the adjustment of pixel size may result in a similar benefit in image fusion accuracy when using hybrid PET/CT scanning.

5.5 Conclusion

We have demonstrated that a small deviation in real image size of PET may occur, as well as a difference in PET image size relative to CT. Although a small deviation in PET image size is not clinically relevant in normal diagnostic procedures, correction of such a difference proves to be beneficial with regard to the accuracy of rigid-body software image fusion. Therefore, it is advisable to re-evaluate PET scanner image size relative to CT images before using high-accuracy rigid-body image fusion with CT.

5.6 References

1. Schoder H, Yeung HW, Gonen M, Kraus D, Larson SM: Head and neck cancer: clinical usefulness and accuracy of PET/CT image fusion. *Radiology*. 2004;231:65-72

2. Schwartz DL, Ford E, Rajendran J, Yueh B, Coltrera MD, Virgin J, Anzai Y, Haynor D, Lewellyn B, Mattes D, Meyer J, Phillips M, Leblanc M, Kinahan P, Krohn K, Eary J, Laramore GE: FDG-PET/CT imaging for preradiotherapy staging of head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys*. 2005;61:129-136

3. Goerres GW, von Schulthess GK, Steinert HC: Why most PET of lung and head-and-neck cancer will be PET/CT. *J Nucl Med*. 2004;45 Suppl 1:66S-71S

4. Scarfone C, Lavely WC, Cmelak AJ, Delbeke D, Martin WH, Billheimer D, Hallahan DE: Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. *J Nucl Med*. 2004;45:543-552

5. Solberg TD, Agazaryan N, Goss BW, Dahlbom M, Lee SP: A feasibility study of 18F-fluorodeoxyglucose positron emission tomography targeting and simultaneous integrated boost for intensity-modulated radiosurgery and radiotherapy. *J Neurosurg*. 2004;101 Suppl 3:381-389

6. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, Corstens FH: PET/CT: panacea, redundancy, or something in between? *J Nucl Med*. 2004;45 Suppl 1:15S-24S

7. Low DA, Mutic S, Dempsey JF, Gerber RL, Bosch WR, Perez CA, Purdy JA: Quantitative dosimetric verification of an IMRT planning and delivery system. *Radiother Oncol*. 1998;49:305-316

8. Lavely WC, Scarfone C, Cevikalp H, Li R, Byrne DW, Cmelak AJ, Dawant B, Price RR, Hallahan DE, Fitzpatrick JM: Phantom validation of coregistration of PET and CT for image-guided radiotherapy. *Med Phys*. 2004;31:1083-1092

9. Hudson HM, Larkin RS: Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imag*. 1994;13:601-609

10. Schroeder W. The Visualization Toolkit: An object-oriented approach to 3D graphics, 3rd edn. New York: Kitware Inc. 2003.

11. Maintz JB, Viergever MA: A survey of medical image registration. *Med Image Anal*. 1998;2:1-3

6

Validated image fusion of PET and CT for external beam radiation therapy in the head and neck area

Wouter V. Vogel [1]
Dominic A.X. Schinagl [2]
Jorn A. van Dalen [1]
Johannes H.A.M. Kaanders [2]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1] and Radiation oncology [2],
Radboud University Nijmegen Medical Centre, the Netherlands

Quarterly Journal of Nuclear Medicine and Molecular Imaging, 2006
In press

Abstract

Integration of PET information into CT-based IMRT allows adaptation of the target volume to functional parameters, but only when the image registration procedure is reliable. The aim of this study was to select the optimal method for software fusion of dedicated PET and CT, and to validate the procedure for IMRT in the head-neck area.

Methods: 15 patients with HNSCC underwent separate CT and FDG-PET, both in a custom-moulded rigid mask fitted with 4 multimodality fiducial markers. Five image registration methods were applied. PET emission and CT were registered manually (ME), and using the landmarks (LM). PET transmission and CT were registered manually (MT), using a mutual information-based method (MI) and using an iterative closest point method (ICP). The error in image registration using each of the methods was determined by evaluation of the markers.

Results: LM showed an excellent average registration error of 1.4 mm at the location of the markers, and 0.3 mm in the planning area. However, this method proved to be laborious. Apart from LM the best method was ICP, with registration errors of 3.0 and 2.0 mm, respectively. For ME the respective errors were 4.7 and 3.5 mm, for MT 3.6 and 2.7 mm, and for MI 5.3 and 4.1 mm.

Conclusions: Image fusion of dedicated PET and CT of the head-neck area can be performed reliably and accurate with no need for laborious markers, using the operator-independent ICP method. The achieved accuracy permits implementation of dedicated PET images in external beam radiation therapy.

6.1 Introduction

External beam radiation therapy of head and neck squamous cell carcinoma (HNSCC) demands accurate dose delivery. Nowadays this can best be achieved with intensity modulated external beam radiation therapy (IMRT), with a good effective spatial accuracy in dose delivery and with high achievable dose gradients (1). As a next step, integration of functional and anatomical information by fusion of PET and CT images allows consideration of biological tumor characteristics in the determination of the IMRT target volume (2-4).

For IMRT, accurate detection and localization of tumor sites is essential. Primary tumors in the head and neck area are often small at the time of discovery, and lymph node metastases tend to be small in size and multiple in number (5). Computed tomography (CT) imaging provides the anatomical reference and electron density information that is mandatory for 3-dimensional planning in IMRT, and can visualize structural information conveniently. Major drawbacks of CT are a low sensitivity for small lymph node metastases and low specificity in marginally enlarged lymph nodes or atypical lesions (6). The inability of CT imaging to differentiate tissue characteristics contributes to this problem, despite the application of intravenous contrast.

Additional functional and molecular information can be provided with positron emission tomography (PET). The radiopharmaceutical ^{18}F-fluor-deoxy-glucose (FDG) quantitatively visualizes glucose metabolism, thus providing a tool for discrimination of normal and malignant tissues. The clinical value of FDG-PET in staging of malignancy in the head and neck area has been demonstrated by some (6-9), although others have reported less impressive results (10). Radiotherapy planning may benefit from improved tumor detection, and from quantitative evaluation of intra-tumoral variations in metabolic activity such as glucose metabolism. Examples of other evaluable biological parameters include hypoxia and proliferation, as visualized with ^{18}F-Misonidazole (FMISO) and ^{18}F-fluor-deoxy-thymidine (FLT), respectively.

Using IMRT, a high dose can be delivered accurately to known tumor sites (11,12). Perhaps equally important, a significant dose reduction to adjacent non-tumor sites can be achieved, resulting in fewer complications and side-effects such as central nervous system damage, mucositis, and loss of parotid gland function (13-15). Furthermore, the 3-dimensional approach of IMRT facilitates intra-tumoral variations in dose delivery, tailored to (regional) specific functional and molecular tumor characteristics. Planning and application of such highly optimized IMRT procedures depends on precise tumor imaging. Any error in the anatomical registration of CT and PET images may lead to erroneous localization or interpretation of lesions, which may subsequently result in suboptimal radiation treatment.

Currently, many perform image fusion of CT and PET for planning of external beam radiation therapy using their own preferred approach to image registration. Some use an integrated PET/CT scanner, others apply software fusion of dedicated CT and PET. Different methods of image registration are available for the latter approach, each with their specific characteristics, and with variable accuracy of image registration. The relative accuracy of the available methods is currently not known. We have hypothesized that for successful application of image fusion in external beam radiation therapy planning – especially for IMRT – the error in image registration should not exceed existing limits in spatial accuracy of image quality and dose delivery. In this paper we evaluate different approaches to high accuracy software PET/CT image fusion, and validate the selected optimal technique for application in IMRT procedures in the head and neck area.

6.2 Methods

A total of 15 patients (mean age 59 years, range 49-74 years) referred for external beam radiation therapy for newly diagnosed HNSCC were included. Twelve patients had carcinoma of the larynx, 3 had carcinoma of the tongue. PET and CT scans were acquired within a maximum interval of one week.

Imaging

CT scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters were 130 kV, 120 mAs, slice distance and slice thickness 3 mm. Images were acquired from the top of the lungs to the base of the skull, with intravenous contrast.

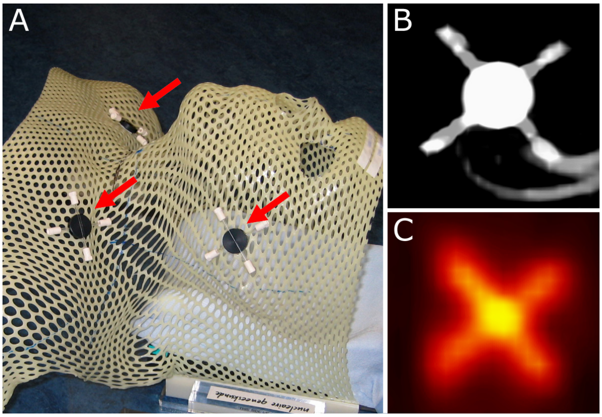


Figure 6.1
Example of the multi-modality markers used for landmark registration and image fusion accuracy evaluation. (A) Markers (red arrows) as placed on the mask. (B) Marker as seen on CT filled with iodine-containing contrast. (C) Marker as seen on PET filled with FDG solution. The center of the markers could be determined with an inter-operator variability well below 1 mm on both CT and PET images.

FDG-PET was acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded. However, glucose levels had to be appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection). A 3D emission scan of the head and neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 minutes after intravenous injection of 250 MBq FDG (Mallinckrodt Medial, Petten, The Netherlands). The acquisition time per bed position was 5 minutes for emission and 3 minutes for the Germanium-based transmission, resulting in a total scanning time of 16 minutes for the two bed positions. Emission and transmission scans were reconstructed using a 2D ordered subset expectation maximization (OSEM) iterative algorithm, with parameters optimized for low photon attenuation in the head and neck area as described elsewhere (16). As during previous studies, a structural difference in real image size between the CT and PET devices of 2.0% in the transversal direction and 0.8% in the axial direction was observed (17), this was corrected by applying a scaling factor.

Patient positioning

During all imaging procedures, patients were placed in radiotherapy position within a custom-moulded rigid mask covering the head, neck and shoulders. Maximum reproducibility in positioning was assured by the use of additional support systems: a flat scanning bed, a customized head support cushion, an intra-oral mould for positioning of the tongue when needed, a standard cushion supporting the knees, and a laser positioning system.

All patients were scanned with 4 multi-modality crosshair fiducial markers firmly attached to the fixation mask in a rectangular configuration around the center of the IMRT planning area, halfway in the anterior-posterior plane. The crosshair markers consisted of two 5 cm glass capillaries positioned in a 90° angle. The capillaries were filled with either 300 mg/ml iodine-containing contrast for CT or 1 MBq/ml FDG for PET. Figure 6.1 shows an example of a mask with fiducial markers in place.

Image registration

The image registration procedure was performed on a PC with in-house developed image viewing and registration software based on the visualization toolkit VTK (18) and the insight segmentation and registration toolkit ITK (19). The software allows rigid-body image registration, i.e. based on 3 translation and 3 rotation parameters. Anatomical image registration of PET images to CT images was performed using five different methods:

1. <u>Method ME</u>: Manual registration of PET emission images to CT. Anatomically correct registration in the field of interest was performed by an experienced investigator, through manual adjustment of the 6 free parameters. The operator used an interface

where 3 variable sections of orthogonal planes (transverse, coronal and sagittal) through the 2 images were displayed simultaneously. Hence, by performing registration in only one section at the time, the 3D registration problem was reduced to a series of 2D problems. The method is therefore fast and easy, but may suffer from operator-dependancy.

2. Method LM: Landmark-based registration of PET emission images to CT. Registration was optimized by manual identification of the centers of the multimodality markers, followed by a minimization of $\sum_{(1-4)}((r_i - T*r'_i)^{\wedge}2)/N$, where r_i are the coordinates of CT landmark i, r'_i are the coordinates of PET landmark i, T is the transformation, and $N=4$ for the number of markers. This minimization provides (by taking the square-root of this minimization result) the fiducial registration error (FRE) (20). This method is subject to operator-dependency due to the manual localization of the markers.
3. Method MT: Manual registration of PET transmission images to CT. Parallel to method ME, anatomically correct registration in the field of interest was pursued through manual adjustment of the 6 free parameters. At the end of the procedure the transmission images were substituted with the emission images. This method is also fast and easy, but may again suffer from operator-dependancy.
4. Method MI: Mutual information based registration of PET transmission images to CT. This method optimizes a functional measuring the similarity of all geometrically corresponding voxel pairs for some feature. The mutual information metric implementation follows the method as specified by Voila and Wells (21,22). In this implementation, probability densities are estimated from the image data using the Parzen-Window scheme (23). This method is available in ITK (19). Its main advantages are that it can work directly with image data as no pre-processing or segmentation is needed. Furthermore, it has an efficient implementation based on stochastic approximation. The parameters of this method (19) have been tuned to our application: the Parzen window width was set to 2, the number of samples to 50, the learning rate to 0.0005, the translation scale to 100 and the number of iterations to 20,000. This provided robust results. At the end of the registration procedure the transmission images were substituted with the emission images. This method is not operator dependant, but does require some computation time.
5. Method ICP: Iterative closest point registration, using surface models of the body contours acquired from PET transmission and CT. A thresholding technique was used on both image sets to create triangulated iso-surfaces of the body, using the VTK method vtkContourFilter (18). Normals were computed for these surfaces for each of the polygonal facets, using the VTK method vtkPolyDataNormals (18). Using these normals, a faceted shading of the surface was obtained that was used as an input for the ICP algorithm. The ICP algorithm (19) has three stages and iterates. Firstly,

random points on the first model (of the transmission PET) are associated to random points on the other model (of the CT) by the nearest neighbor criterion. Secondly, transformation parameters are estimated using a mean square cost function. Next, the transformation is applied. Then, the iteration procedure starts by re-associating points, etc. The procedure continues untill convergence has been reached. In this study 5000 random points were used for each model and the convergence criterion was to stop when the change between two successive iterations fell below a threshhold of 0.001 mm, or when the number of iterations exceeded 3000. At the end of the procedure the transmission PET images were substituted with the emission images. The ICP method is illustrated in figure 6.2. This method is not operator dependant, but does require some computation time.

For the automatic registration methods MI and ICP the algorithm input was restricted to a volume of interest by defining a 3-dimensional box, containing the head and neck area, but excluding bodyparts that tend to keep freedom of movement within the mask (i.e. the shoulders). With the exception of the landmark-based method LM, the fiducial markers were not considered during the image registration process. In the PET transmission images the markers were not visible, and all representations of the markers were removed from the emission images prior to image registration.

Assessment of image registration

The fiducial markers were used for assessment of the accuracy in image registration. The location of the center of each marker was determined on both CT and PET emission images. The inter-operator variation of the manual localization of the center of the markers was evaluated by analysis of 12 markers in a separate session by two operators. The difference in position of corresponding markers on CT and PET was determined, representing the error in image registration at the location of the markers. Subsequently, for each patient the location of a hypothetical marker in the center of the IMRT planning area was determined, by calculating the geometrical center of all surrounding markers. The difference in position of this hypothetical marker on CT and PET was determined to estimate the image registration error in the region of the IMRT planning area.

Statistical analysis

Differences in the image registration errors for the applied methods were evaluated using a repeated measures ANOVA with Bonferroni’s correction for multiple comparisons, using GraphPad InStat version 3 (GraphPad Software, San Diego CA, USA). The level of significance was set at 0.05.

6.3 Results

The inter-operator variability in determination of the center of the markers was well below 1 mm in both types of scans. For CT images the average difference between the two operators was 0.54 mm (S.D. 0.28 mm). For PET images the average difference was 0.42 mm (S.D. 0.19 mm).

Application of LM achieved the best results (P < 0.001 as compared to all other methods). LM showed an average registration error of 1.4 mm (range 0.3 – 3.8, S.D. 0.8) at the location of the markers, and a calculated average registration error in the center of the planning area of 0.3 mm (range 0.0 – 0.6, S.D. 0.2). The second-best method proved to be ICP, with an average error in image registration of 3.0 mm (range 0.5 – 8.8, S.D. 0.9) at the location of the markers and 2.0 mm (range 0.6 – 4.3, S.D. 1.1) at the center of the planning area. The average error in the center was above the limit of 3 mm in three cases (20%), the largest error in the planning area being 4.3 mm.

For method ME the respective registration errors were 4.7 mm (range 1.2 – 13.6, S.D. 2.2) and 3.5 mm (range 0.7 – 6.3, S.D. 1.6), for method MT 3.6 mm (range 0.5 – 10.6, S.D. 1.5) and 2.7 mm (range 1.1 – 6.1, S.D. 1.6), and for method MI 5.3 mm (range 1.0 – 18.3, S.D. 2.0) and 4.1 mm (range 1.9 – 6.8, S.D. 1.4). As compared to ICP, the methods ME and MI were significantly less accurate in image registration (P < 0.05). MT was not significantly different as compared to method ICP. An example of the results after image registration is shown in figure 6.3. The quantitative results are depicted in figure 6.4.

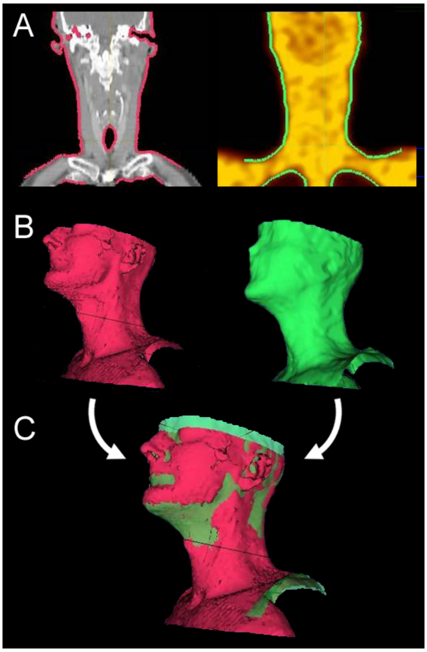
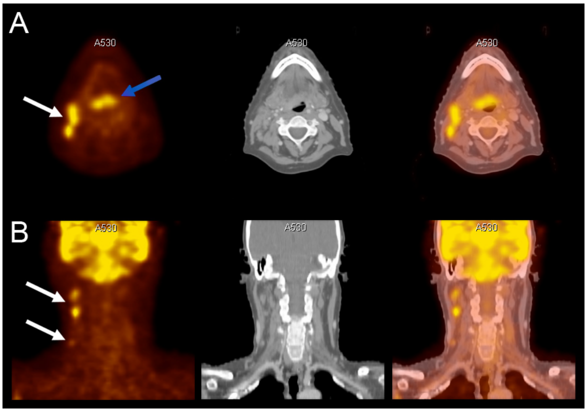


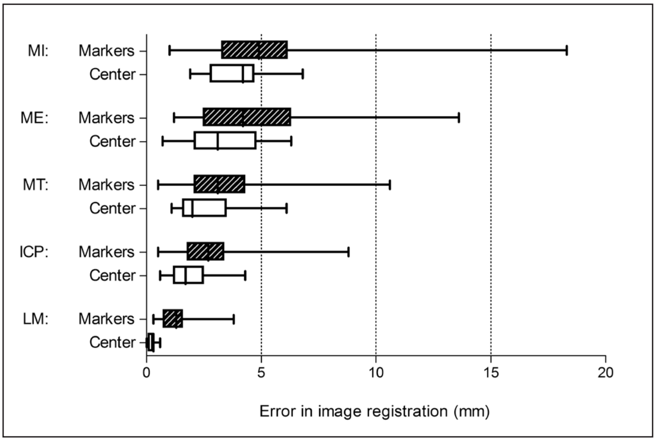
Figure 6.2
Iterative closest point (ICP) registration of PET transmission images to CT. (A) On both CT and the PET transmission scan, a surface model of the body contour is generated using a thresholding technique. (B) Both surface models are represented as a 3-dimensional structure. (C) Automatic image registration is performed by iteratively minimizing the distance between the surface models.

Figure 6.3
Example of image fusion using method ICP. Transverse (A) and coronal (B) sections of FDG-PET (left), CT (middle) and fused PET-CT (right) in a patient with carcinoma at the base of the tongue (blue arrow) and multiple pathological lymph nodes on the right side of the neck (white arrows). The extent of the malignancy was difficult to appreciate on the CT, while fused PET/CT images allowed clear delineation of the primary tumor, as well as identification of pathological lymph nodes.



During the PET image acquisition, small positioning problems occurred in 4 patients, due to relative inexperience of the PET operating personnel with the fixation mask. Examples are a minor backwards tilt of the head within the mask, or an incorrectly attached part of the mask near a shoulder. This resulted in minor visually discernible positioning differences between PET and CT scans, in all cases below 1 cm and mostly well outside the IMRT planning area. All 3 patients who showed a registration error above 3 mm using method ICP were subject to such errors.

Figure 6.4
Accuracy of image registration. The errors in image registration, as performed by the different applied techniques (MI = mutual information, ME = manual emission, MT = manual transmission, ICP = iterative closest point, LM = landmark). The graph shows that the lowest error in the center of the IMRT field can be achieved using the landmark-based method (LM).



6.4 Discussion

Image fusion may be performed with software-based image registration of dedicated PET and CT. However, image registration of dedicated PET and CT has a risk for introducing errors (24). Although high accuracy in image registration has been shown using phantoms (25), a time interval between scans and repeated positioning may result in introduction of differences when imaging patients. Furthermore, the image registration procedure itself is hampered by limited visualization of normal anatomical structures on FDG-PET. The risk for potentially significant errors in PET/CT image registration emphasizes the need for a structured and validated approach.

The spatial resolution of PET imaging is in the range of 4–7 mm (full width half maximum) for most currently available scanners, although the image quality may be better in the head and neck area where photon attenuation is low (16). The geometrical uncertainty in IMRT dose delivery is in the range of 2–3 mm (1). Therefore, an upper limit for the error in image registration of 3 mm can be considered acceptable for application in IMRT planning.

Fiducial markers
At least three fiducial markers are needed to define position and orientation of a rigid object in a 3-dimensional space. In these studies 4 markers were used, in a rectangular configuration in the coronal plane around the planning area to equally represent the whole head/neck area. The anterior and posterior sides of the mask were not marked with separate markers, but these areas were considered not very prone to additional local positioning errors because of rigid attachment to the scanning table and adequate fixation of the nose.

Fiducial markers can also be used for evaluation of the local image registration error. This local error is influenced by non-rigid positioning differences within the mask. The extent of such positioning differences is unclear. According to van Lin et al. the error in repeated positioning in a customized mask is in the range of 3-4 mm (26), but this includes other factors such as visual and manual correlation with laser guides. In this paper, the relatively good results of image registration suggest that on average patient positioning was adequate, and well within the range found by van Lin et al. Nevertheless, some of the patients were indeed subject to slightly suboptimal positioning, due to relative inexperience of PET personnel in postioning of the customized mask. This suggests that the accuracy of the image registration may improve even further, as experience grows.

In this study, the errors in image registration were measured by markers placed well outside the target volume for radiotherapy, which can be considered a ‘worst case scenario’. The much more relevant error in the center of the planning area was estimated by the error in a mathematically derived hypothetical marker, defined by the geometrical center of all available surrounding markers. A drawback of this method is that it represents a theoretical value. However, there is no reliable direct method to evaluate the errors in image registration within the neck of a patient, due to the lack of anatomical landmarks that are well demarcated both on CT and PET transmission images, the unfeasibility of marker placement in such areas, and unavoidable physiologic processes such as swallowing. Wong et al. have described anatomical landmarks that may be identified in the head and neck area on both PET and CT (27), but we considered the visibility of such landmarks very poor, and also considered the suggested accuracy of localization of 1.3 – 8.2 mm insufficient. Furthermore, the artificial markers were chosen as these do not depend on tracer biodistribution and uptake, and as manual localization will be less operator dependent. Therefore, the calculated error in the center of the planning area, derived from fiducial markers at the edges, was considered the best available indicator of the real error in image registration.

Fitzpatrick et al. have previously published a method for quantitative evaluation of registration errors using recognisable landmarks (28,29). The so-called "Target Registration Error" (TRE) expresses the displacement between any two corresponding points in the images, in relation to the fiducial registration error (FRE) of available landmarks. This TRE is considered to be the most accurate mathematical representation of the real registration error at a specific point. Given the configuration and number of the markers in our study, the calculated average position of the hypothetical marker in the center of the planning area approaches the local TRE as defined by Fitzpatrick. The error in areas further away from the planning area (e.g. at the anterior and posterior borders of the head and neck area) will be somewhat underestimated, but this is considered not very relevant for treatment planning for areas in the center field of view.

Image registration

The best results in image registration accuracy were achieved by the landmark-based method. These values were considered a 'gold standard' in rigid-body image registration accuracy as they reflect the mathematically achievable accuracy, including all unavoidable positioning differences, mask deformations, patient motion due to e.g. swallowing, and sampling errors in determining the position of the markers. The average registration error at the location of the landmarks is better than the values observed when using anatomical landmarks by Wong et al., probably because the manual localization procedure of artificial markers is more accurate and less operator independant (27). Furthermore, our results with LM are very similar to those reported with integrated PET/CT scanners, which provide an excellent tool for image fusion for IMRT planning in the head and neck area (30,31).

In practice, LM proved to be relatively laborious and cumbersome, as for each procedure the markers had to be filled manually, placed on the masks and subsequently removed again. This may partially be overcome by using commercially available multimodality markers, containing a solid long lived PET source. However, an adequate method with no need for markers may still be preferable.

The ICP method proved to be the second best method, as compared to the landmark-based registration method. With ICP, the average error in image registration was 2.0 mm in the center of the planning area, which was within the limits of accuracy of IMRT (1). ICP proved to be significantly better than the ME and mutual information based methods. A further advantage of ICP over the manual methods is the operator independency. An advantage over the mutual information method is the independency of initial registration differences between PET and CT. The results of the ICP method in this study were better than previously published results of software PET/CT fusion of the brain in a single patient using a mutual information algorithm by Lavely et al. (25). Obviously, a direct comparison with the head-neck area is not possible. Also, Lavely et al. did show more precise registration in a phantom, but it is not very realistic to extrapolate these results to clinical procedures.
Using ICP, three patients still showed an error of more than 3 mm in image registration in the center of the planning area. In these cases, less experienced personnel positioned the patient in the fixation mask. Thus, it is likely that a smaller registration error of the ICP method (and possibly the other methods) than reported in our series is feasible, when experience in patient positioning increases. This also illustrates that patient positioning could be the limiting factor in image registration, rather than the ICP method itself. This supports the conclusion that PET-CT image registration using the ICP method can be applied in the IMRT procedure.
The method ME performed significantly less accurate than the ICP method. Furthermore, as opposed to the ICP method, both manual registration methods (ME and MT) are operator-dependent, which is a clear disadvantage. Therefore, the manual methods were considered suboptimal for procedures in the head and neck area as compared to ICP.

The performance of the MI based algorithm was somewhat disappointing, as the method performed significantly less accurate than the ICP method, with an average error larger than 3 mm. We have observed that the algorithm consistently converges to a reasonable registration, but tends to deviate slightly in the final registration result as an intelligent evaluation of edges and symmetry seems to be lacking, as opposed to the ICP method. Furthermore, the results may vary with the extent and direction of the registration mismatch that the MI algorithm is confronted with initially. The results of an MI-based algorithm depend on several parameters, such as the number of iterations. We fine-tuned our parameters, but were unable to improve the results to the level of the ICP results. Therefore, we consider the Parzen

Window implementation of the MI method suboptimal for procedures in the head and neck area as compared to ICP.
The ICP method in this study was based on registration of PET transmission images to CT. This implies that PET without transmission images cannot be registered using the ICP method. This is not a relevant problem, as PET-based quantitative tissue characterization for IMRT already requires correction for photon attenuation. A further important advantage of using transmission images is the independence of the choice of the radiopharmaceutical (i.e. FLT, F-Misonidazole). This allows image registration also when delineation of anatomical landmarks in the emission images is relatively poor and tumor uptake is low. Theoretically, image registration based on transmission images may fail in case of patient motion between the acquisition of emission and transmission, but the use of a fixation mask will prevent such problems. The fixation mask itself is visible on CT images and may slightly influence PET transmission images, despite its limited thickness and density. Visualization of the mask on PET transmission images would be homogeneous and symmetrical, and is not considered a problem for image registration. Especially the ICP method is unaffected by symmetrical factors, such as the threshold level on PET images. The evaluation of image registration with fiducial markers is not influenced by the presence of the mask.

The accuracy and robustness as achieved by the observer independent ICP method has not been published before. Only a few studies have evaluated the accuracy of rigid image registration in the head and neck area. As stated above, one study has shown the feasibility of image fusion in 30 patients using manually selected anatomical landmarks (27), although the method is very operator dependent and the average registration error of 3.8 mm (range 1,3 – 8,2 mm) is considered insufficient for high-accuracy procedures such as IMRT. Another study by Klabbers et al. has also demonstrated the feasibility of mutual information based registration with transmission images (32), but the image registration was only evaluated for the object as a whole using a full-circle method, which provides no information about local and systematic registration errors. The claimed average registration error of 4 mm is exclusive additional errors due to rotation, and neglects local errors due to positioning differences. A study by Nishioka et al. used image registration based on the brain only, and used visual evaluation only (33).

Other image registration techniques
We have evaluated 5 different approaches to PET-CT image registration. More techniques are available, as published in an extensive listing by Maintz et al. (20). In this study, a method has been selected from each category of techniques, unless inappropriate (i.e. not applicable for inter-modality registration). Other techniques are not yet established, such as non-rigid (elastic) registration.

The high accuracy of image registration presented in this study is only valid for rigid-body image registration. This limits the applicability of this approach to IMRT planning in the head and neck area, where measures can be taken to ensure reproducible positioning. It is likely that image registration of non-rigid body parts where immobilization and positioning tools are less effective – such as the chest and abdomen – will be less accurate, or will require different approaches to image registration, tailored to the specific situation.

Hybrid PET/CT

For hybrid PET/CT scanning, phantom studies have demonstrated an image registration accuracy in the range of 0-2 mm, and for patients in the range of 1-3 mm in the head and neck area (30). Furthermore, patient repositioning between the acquisition of PET and CT images is no longer required. However, many hospitals do not have access to an integrated PET/CT scanner system. As the clinical demand for functional imaging in radiation therapy planning is high, also in hospitals without a hybrid PET/CT scanner, validation of software image fusion remains of relevance.

6.5 Conclusions

High-accuracy rigid-body image registration of dedicated PET and CT of the head and neck area can be performed adequately and reliably, with or without multi-modality markers. Without using markers, the iterative closest point registration method of CT and PET-transmission images proved to be the most accurate operator-independent image registration method. The achieved accuracy – with an average error lower than 3 mm in image registration – permits implementation of dedicated PET images in IMRT planning and therapy. Our results validate the use of software image fusion of PET and CT for IMRT planning in the head and neck area, thus permitting application of this technique whenever dedicated PET and CT are available.

6.6 References

1. Low DA, Mutic S, Dempsey JF, Gerber RL, Bosch WR, Perez CA, and Purdy JA. Quantitative dosimetric verification of an IMRT planning and delivery system. Radiother. Oncol. 1998;49:305-316
2. Ling CC, Humm J, Larson S, Amols H, Fuks Z, Leibel S, and Koutcher JA. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. Int. J. Radiat. Oncol. Biol. Phys. 2000;47:551-560

3. Scarfone C, Lavely WC, Cmelak AJ, Delbeke D, Martin WH, Billheimer D, and Hallahan DE. Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. J. Nucl. Med. 2004;45:543-552

4. Solberg TD, Agazaryan N, Goss BW, Dahlbom M, and Lee SP. A feasibility study of 18F-fluorodeoxyglucose positron emission tomography targeting and simultaneous integrated boost for intensity-modulated radiosurgery and radiotherapy. J. Neurosurg. 2004;101 Suppl 3:381-389

5. Stuckensen T, Kovacs AF, Adams S, and Baum RP. Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J. Craniomaxillofac. Surg. 2000;28:319-324

6. Bruschini P, Giorgetti A, Bruschini L, Nacci A, Volterrani D, Cosottini M, Ursino F, Mariani G, and Fattori B. Positron emission tomography (PET) in the staging of head neck cancer: comparison between PET and CT. Acta Otorhinolaryngol. Ital. 2003;23:446-453

7. Braams JW, Pruim J, Freling NJ, Nikkels PG, Roodenburg JL, Boering G, Vaalburg W, and Vermey A. Detection of lymph node metastases of squamous-cell cancer of the head and neck with FDG-PET and MRI. J. Nucl. Med. 1995;36:211-216

8. Laubenbacher C, Saumweber D, Wagner-Manslau C, Kau RJ, Herz M, Avril N, Ziegler S, Kruschke C, Arnold W, and Schwaiger M. Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous-cell carcinomas. J. Nucl. Med. 1995;36:1747-1757

9. Teknos TN, Rosenthal EL, Lee D, Taylor R, and Marn CS. Positron emission tomography in the evaluation of stage III and IV head and neck cancer. Head Neck. 2001;23:1056-1060

10. Stoeckli SJ, Steinert H, Pfaltz M, and Schmid S. Is there a role for positron emission tomography with 18F-fluorodeoxyglucose in the initial staging of nodal negative oral and oropharyngeal squamous cell carcinoma. Head Neck. 2002;24:345-349

11. De Neve W and Duthoy W. Intensity-modulated radiation therapy for head and neck cancer. Expert. Rev. Anticancer Ther. 2004;4:425-434

12. Zhou J, Fei D, and Wu Q. Potential of intensity-modulated radiotherapy to escalate doses to head-and-neck cancers: what is the maximal dose? Int. J. Radiat. Oncol. Biol. Phys. 2003;57:673-682

13. Lin A, Kim HM, Terrell JE, Dawson LA, Ship JA, and Eisbruch A. Quality of life after parotid-sparing IMRT for head-and-neck cancer: a prospective longitudinal study. Int. J. Radiat. Oncol. Biol. Phys. 2003;57:61-70

14. Parliament MB, Scrimger RA, Anderson SG, Kurien EC, Thompson HK, Field GC, and Hanson J. Preservation of oral health-related quality of life and salivary flow rates after inverse-planned intensity- modulated radiotherapy (IMRT) for head-and-neck cancer. Int. J. Radiat. Oncol. Biol. Phys. 2004;58:663-673

15. Cozzi L, Fogliata A, Bolsi A, Nicolini G, and Bernier J. Three-dimensional conformal vs. intensity-modulated radiotherapy in head-and-neck cancer patients: comparative analysis of dosimetric and technical parameters. Int. J. Radiat. Oncol. Biol. Phys. 2004;58:617-624

16. Vogel WV, Wensing BM, van Dalen JA, Krabbe PF, van den Hoogen FJ, and Oyen WJ. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. Eur. J. Nucl. Med. Mol. Imaging. 2005;32:1276-1282

17. Vogel WV, van Dalen JA, Schinagl DAX, Kaanders JH, Huisman H, Corstens FH, and Oyen WJ. Correction of an image size difference between PET and CT improves image fusion of dedicated PET and CT. Submitted. 2006;

18. Schroeder W. The Visualization Toolkit: An object-oriented approach to 3D graphics, 3rd edn. New York: Kitware Inc. 2003;

19. Ibanez L, Schroeder W, Ng L, and Cates J. The ITK Software Guide: The Insight Segmentation and Registration Toolkit, version 1.4. New York: Kitware, Inc. 2003;

20. Maintz JB and Viergever MA. A survey of medical image registration. Med. Image Anal. 1998;2:1-36

21. Viola P and Wells WM. Alignment by maximization of mutual information. Proc. 5th Int. Conf. on Computer Vision (Boston) (Cambridge MA: MIT Press). 1995;16-23

22. Viola P and Wells WM. Alignment by maximization of mutual information. Int. J. of Computer Vision. 1997;137-154

23. Duda and Hart, P. Pattern Classification and Scene Analysis. (New York: Wiley).

24. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, and Corstens FH. PET/CT: panacea, redundancy, or something in between? J. Nucl. Med. 2004;45 Suppl 1:15S-24S

25. Lavely WC, Scarfone C, Cevikalp H, Li R, Byrne DW, Cmelak AJ, Dawant B, Price RR, Hallahan DE, and Fitzpatrick JM. Phantom validation of coregistration of PET and CT for image-guided radiotherapy. Med. Phys. 2004;31:1083-1092

26. van Lin EN, van d, V, Huizenga H, Kaanders JH, and Visser AG. Set-up improvement in head and neck radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support. Radiother. Oncol. 2003;68:137-148

27. Wong WL, Hussain K, Chevretton E, Hawkes DJ, Baddeley H, Maisey M, and McGurk M. Validation and clinical application of computer-combined computed tomography and positron emission tomography with 2-[18F]fluoro-2-deoxy-D-glucose head and neck images. Am. J. Surg. 1996;172:628-632

28. Fitzpatrick JM, West JB, and Maurer CR, Jr.. Predicting error in rigid-body point-based registration. IEEE Trans. Med. Imaging. 1998;17:694-702

29. Fitzpatrick, Hill, D. L. G., and Maurer, C. R. J. Image registration. Handbook of Medical Imaging (Bellingham, Washinton: SPIE Press):447-513.

30. Ciernik IF, Dizendorf E, Baumert BG, Reiner B, Burger C, Davis JB, Lutolf UM, Steinert HC, and von Schulthess GK. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. Int. J. Radiat. Oncol. Biol. Phys. 2003;57:853-863

31. Heron DE, Andrade RS, Flickinger J, Johnson J, Agarwala SS, Wu A, Kalnicki S, and Avril N. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. Int. J. Radiat. Oncol. Biol. Phys. 2004;60:1419-1424

32. Klabbers BM, de Munck JC, Slotman BJ, Langendijk HA, de BR, Hoekstra OS, Boellaard R, and Lammertsma AA. Matching PET and CT scans of the head and neck area: development of method and validation. Med. Phys. 2002;29:2230-2238

33. Nishioka T, Shiga T, Shirato H, Tsukamoto E, Tsuchiya K, Kato T, Ohmori K, Yamazaki A, Aoyama H, Hashimoto S, Chang TC, and Miyasaka K. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. Int. J. Radiat. Oncol. Biol. Phys. 2002;53:1051-1057

7

Comparison of 5 segmentation tools for FDG-PET based target volume definition in head and neck cancer

Dominic A.X. Schinagl [1]
Wouter V. Vogel [2]
Aswin L. Hoffmann [1]
Jorn A. van Dalen [2]
Wim J.G. Oyen [2]
Johannes H.A.M. Kaanders [1]

Departments of Radiation oncology [1] and Nuclear medicine [2],
Radboud University Nijmegen Medical Centre, the Netherlands

International Journal of Radiation Oncology, Biology, Physics
In press.

Abstract

Target volume delineation for radiation treatment in the head and neck area is traditionally based on physical examination, CT, and MRI. Additional molecular imaging with FDG-PET may improve definition of the gross tumor volume (GTV), but there is no consensus on the method of tumor delineation. In this study, five methods for tumor delineation on FDG-PET are compared with CT-based delineation.

Methods: Seventy-eight patients with stage II-IV squamous cell carcinoma of the head and neck area underwent co-registered CT and FDG-PET. The primary tumor was delineated on CT (GTV_{CT}), and five PET-based GTVs were obtained: visual interpretation (GTV_{VIS}), applying an isocontour of a standardized uptake value of 2.5 (GTV_{SUV}), using a fixed threshold of 40% and 50% of the maximum signal intensity ($GTV_{40\%}$ and $GTV_{50\%}$), and applying an adaptive threshold based on the signal to background ratio (GTV_{SBR}). Absolute GTV volumes were compared, and overlap analyses were performed.

Results: The GTV_{SUV} method failed to provide successful delineation in 45% of cases. For the other PET delineation methods, the volume and shape of the GTV were heavily influenced by the choice of the segmentation tool. On average, all threshold-based PET-GTVs were smaller than on CT. Nevertheless, PET frequently detected significant tumor extension outside the GTV_{CT} (15-34% of PET-volume).

Conclusions: The choice of a segmentation tool for target volume definition of head and neck cancer based on FDG-PET images is not trivial, as it influences both volume and shape of the resulting GTV. With adequate delineation, PET may add significantly to CT and physical examination-based GTV definition.

7.1 Introduction

Progress in radiation oncology enables delivery of radiation treatment with increasing geometric precision. This requires a re-evaluation of target volume delineation, which is traditionally based on physical examination, computed tomography (CT) and magnetic resonance imaging (MRI). In recent years, new methods have been introduced for visualization of tumor tissue. In addition to anatomical data as supplied by CT and MRI, 'functional' and 'molecular' imaging techniques, such as positron emission tomography (PET), single-photon emission computed tomography (SPECT) and magnetic resonance spectroscopy (MRS), allow visualization of biological characteristics, with several potential advances. The primary tumor may be identified more accurately, with consequences for the size and shape of the gross tumor volume (GTV). Tumor characteristics relevant for radiation sensitivity can be visualized (e.g. hypoxia), which may assist in the selection of patients for customized treatments (1). Also, intra-tumoral heterogeneity of these characteristics may be identified, providing an opportunity for 'dose painting' (2). Finally, when imaging modalities become more accurate, the inter- and intra-observer variations in tumor delineation will decrease, resulting in improved standard of care.

Metabolic information, as provided by imaging ^{18}F-fluor-deoxy-glucose (FDG) with PET, has been incorporated into target volume delineation by many groups (3). Tumor localizations can be identified and localized with high sensitivity, due to the high contrast resolution of PET. However, application of FDG-PET data for target volume delineation is not straightforward, as identification of tumor boundaries on PET suffers from a relatively low spatial resolution and a 'blurry' appearance of lesions. Furthermore, FDG-PET is usually interpreted qualitatively, whilst in radiation oncology a more quantitative approach is required for tumor contouring (4). Currently, various methods for FDG-PET based target volume definition are in use. Visual interpretation is the most commonly used method (5-12). This method, however, is susceptible to the window-level settings of the images and is highly operator dependent. Therefore, more objective methods have been explored. Examples are isocontouring based on either a standardized uptake value (SUV) of 2.5 (10,13,14), a fixed threshold of the maximum signal intensity (15-21), or on a threshold which is adaptive to the signal to background ratio (22). The utility of these methods for tumor delineation in the head and neck area is currently unknown.

The choice of method for tumor delineation on FDG-PET may influence GTV determination, with consequences for the outcome of radiation therapy. The aim of this study was to compare different methods for tumor delineation with FDG-PET, relative to CT-based delineation, for radiation therapy planning in head and neck cancer patients.

7.2 Methods and materials

Patients

Seventy-eight patients (59 males and 19 females, median age 61 years, range 43-86 years) with stage II-IV squamous cell carcinoma of the head and neck area, eligible for primary curative radiotherapy, were prospectively enrolled from June 2003 until July 2006. The tumor characteristics are summarized in table 7.1. The study was approved by the Ethics Committee of the Radboud University Nijmegen Medical Centre and all patients provided informed consent.

Image acquisition

Prior to treatment, a CT scan and an FDG-PET scan were acquired in radiation treatment position, the patient being immobilized by a custom-made rigid mask covering head, neck and shoulders. Maximum reproducibility in positioning was assured by the use of additional support systems: a flat scanning bed, a customized head support cushion, an intra-oral mould when indicated, a standard cushion supporting the knees, and a laser positioning system. The median interval between CT and FDG-PET was 3 days (range 0-10 days). The CT scan was always performed prior to the PET-scan. CT scans were acquired using a multislice spiral CT scanner (Philips AcQsim, Philips, Cleveland, USA). Scanning parameters included 130 kV, 120 mAs, slice distance and slice thickness 3 mm, scanning from above the frontal sinuses to below the clavicles, with intravenous contrast.

FDG-PET was acquired using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients with diabetes mellitus were not excluded, however, glucose levels had to be appropriately regulated (glucose levels at time of FDG injection < 10 mmol/l, no insulin administration prior to FDG injection). A 3D emission scan of the head and neck area and a 2D Germanium-68 based transmission scan for attenuation correction were acquired 60 minutes (median 64 minutes, S.D. ± 11.4) after intravenous injection of 250 MBq FDG (Mallinckrodt Medical, Petten, the Netherlands). The acquisition

Patient and tumor characteristics	
Tumor site	
Oral Cavity	6
Oropharynx	31
Hypopharynx	9
Larynx	32
T-stage	
T1	1
T2	16
T3	39
T4	22

Table 7.1
Tumor characteristics.

time per bed position was 5 minutes for emission and 3 minutes for the Germanium-based transmission scan, resulting in a total scanning time of 16 minutes for two bed positions. Emission and transmission scans were reconstructed using a 2D ordered subset expectation maximization (OSEM) iterative algorithm, with parameters (4 iterations, 16 subsets) optimized for low photon attenuation in the head and neck area as described elsewhere (23).

Image registration

The image registration procedure has been described in detail previously (24). In brief, three-dimensional surface models were automatically derived from both the CT and the PET transmission images. These models were anatomically co-registered using an operator-independent iterative closest point algorithm, with an average registration error of 2.0 mm at the centre of the planning area. Afterwards, the PET transmission images were replaced with the PET emission images. In addition, a second PET data set was generated in which the original values were replaced with calculated SUV values. SUV was defined as the voxel value of detected activity (in Bq/ml) multiplied by the weight of the patient (in Kg) divided by the activity at the beginning of the scan (in Bq) multiplied by 1000 (25).

The CT and the two PET data sets were transferred via DICOM to a Pinnacle^3 treatment planning system (Philips Medical Systems, Andover, MA, USA) for target volume definition and subsequent volume analysis.

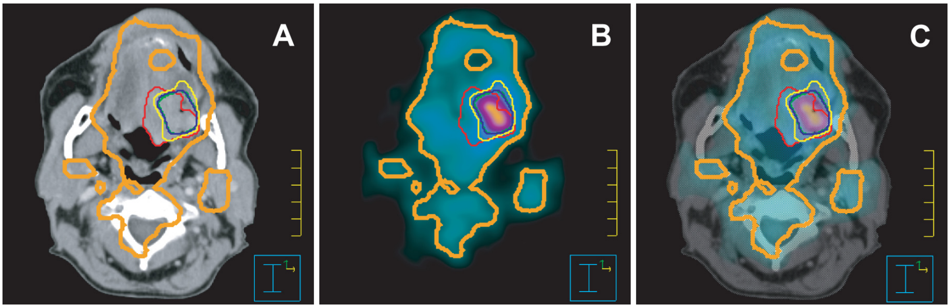


Figure 7.1
CT scan (A), corresponding FDG-PET scan (B) and fused image (C) of a patient with a T4N2M0 tongue carcinoma, showing differences in target volume definition. Indicated are GTV_{CT} *(red),* GTV_{VIS} *(light green),* GTV_{SUV} *(orange),* $GTV_{40\%}$ *(yellow),* $GTV_{50\%}$ *(blue), and* GTV_{SBR} *(dark green).* GTV_{SUV} *is unsuccessful in this case, due to inclusion of large areas of normal background tissue. Note that all other PET-based delineations indicate greater tumor extension towards lateral side and less towards medial side compared to CT delineation. Also note that on this transversal slice* $GTV_{50\%}$ *and* GTV_{SBR} *are indistinguishable.*

Target volume definition

The primary tumor was delineated on CT and FDG-PET images by two experienced radiation oncologists in consensus. The role of FDG-PET in detection of metastatic lymph nodes will be the subject of a separate analysis.

On CT images, manual delineation of the GTV_{CT} was performed according to current clinical protocols, using information gathered from physical examination, available diagnostic workup imaging modalities (CT and/or MRI, examination under general anesthesia) and the CT in treatment position. When the radiation oncologists were drawing the GTV_{CT} contours, the FDG-PET images were blinded.

Five PET-based GTVs were obtained using different delineation approaches. Visual delineation (GTV_{VIS}) was performed by contouring FDG activity clearly above normal background activity. Localizations with increased FDG uptake were classified malignant in consensus with an experienced nuclear medicine physician. The other (threshold based) GTVs were obtained using in-house developed software scripts for the $Pinnacle^3$ treatment planning system. SUV-based delineation was obtained by applying an isocontour of SUV=2.5 (GTV_{SUV}) around the tumor. Two thresholds were based on fixed percentages of the maximum signal intensity in the primary tumor, of 40% ($GTV_{40\%}$) and 50% ($GTV_{50\%}$) respectively. Finally, an adaptive threshold delineation (GTV_{SBR}) based on the signal-to-background ratio (SBR) was performed, as developed at Université St. Luc in Brussels, Belgium (22). The maximum signal intensity was defined as the mean activity of the hottest voxel and its eight surrounding voxels in a transversal slice. The mean background activity was obtained in a manually defined region of interest of approximately 10 cm^3 in the left neck musculature, far away from the primary tumor and any involved lymph nodes. Prior to delineation, scanner-specific variables that were needed for calculation of the GTV_{SBR} were derived by a phantom experiment as described by Daisne et al. (22). In brief, hot spheres with different sizes in a 6.5L Jaszczak phantom were imaged at different image contrast ratios. Optimal delineation thresholds were determined by minimizing the square difference between true and measured sphere volumes. These results were used to find the parameters a and b in the algorithm: Threshold=a+b×1/SBR (in our setting a=44.1, and b=70.4). These data were consistent in replicate experiments.

For all FDG-PET delineation methods, air cavities were excluded from the GTV. This was achieved using automatic contouring of air cavities on CT (Hounsfields units 0-900) and subsequent subtraction from the GTV using an in-house developed $Pinnacle^3$ script. Results obtained by automated delineation algorithms were checked visually before acceptation. A delineation was considered unsuccessful if the resulting GTV included significant volumes of tissue that were clearly normal at visual interpretation.

Volume analysis

Absolute volumes of different GTVs were compared using a paired t-test. In addition, for all PET GTVs, in-house developed scripts were used to calculate: 1. the overlap volume of GTV_{CT} and GTV_{PET}, where overlap was expressed as the overlap volume of GTV_{CT} and GTV_{PET} relative to GTV_{CT} (overlap $fraction_{CT}$ (OF_{CT})), and as the overlap volume of GTV_{CT} and GTV_{PET} relative to the five PET-based GTVs (overlap $fraction_{PET}$ (OF_{PET})); 2. the volume enclosed by GTV_{CT} but not by GTV_{PET} relative to GTV_{CT}, which is $1 \cap OF_{CT}$; and 3. the volume enclosed by GTV_{PET} but not by GTV_{CT} relative to GTV_{PET}, which is $1 \cap OF_{PET}$. Correlations between overlap fractions were assessed by linear regression analysis.

7.3 Results

Seventy-eight patients were included in this study. Of these, 77 data sets were available for analysis; one patient was excluded as the primary tumor, a T2N2cM0 oropharyngeal carcinoma, was not visualized by FDG-PET.

The GTV_{VIS} could be generated for all 77 patients. The GTV_{SBR} segmentation tool resulted in an unsuccessful volume definition in two patients. This was observed in four patients for both the $GTV_{40\%}$ and the $GTV_{50\%}$, two of whom also had an unsatisfactory GTV_{SBR}. The GTV_{SUV} determination was not successful in 35 patients, including the four patients mentioned above. As a consequence, this latter method was not further evaluated. Unsuccessful delineation was not correlated with specific tumor subsites or T-stages. All unsatisfactory volumes were largely oversized, being at least 300 cm^3. An example of an inadequate GTV_{SUV} is depicted in figure 7.1.

GTV volume and overlap fractions for various segmentation tools								
	mean absolute volume		mean OF_{CT}			mean OF_{PET}		
	cm^3	95% CI	cm^3	95% CI		cm^3	95% CI	
GTV_{CT}	22.7	[17.4 - 27.9]	-			-		
GTV_{VIS}	21.5	[16.5 - 26.6]	0.61	[0.56 - 0.66]		0.66	[0.66 - 0.70]	
$GTV_{40\%}$	16.4	[13.2 - 19.6]	0.55	[0.51 - 0.58]	*	0.72	[0.67 - 0.77]	°
$GTV_{50\%}$	10.5	[8.2 - 12.7]	0.39	[0.36 - 0.43]	**	0.80	[0.76 - 0.85]	°°
GTV_{UCL}	11.2	[8.2 - 12.9]	0.43	[0.35 - 0.51]	***	0.85	[0.74 - 0.97]	°°°

Table 7.2

GTV volume and overlap fractions for various segmentation tools. CI: Confidence Interval. OF_{CT}: $GTV_{CT} \cap GTV_{PET} / GTV_{CT}$. OF_{PET}: $GTV_{CT} \cap GTV_{PET} / GTV_{PET}$. *p-values (relative to* GTV_{VIS}*): * p = 0.0007, ** p < 0.0001, *** p < 0.0001, ° p = 0.01, °° p < 0.0001, °°° p = 0.001.*

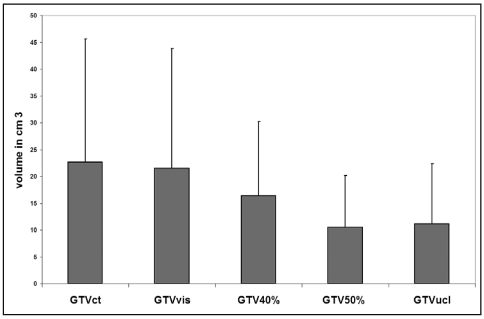


Figure 7.2
Average absolute volumes of the various GTV methods. Error bars indicate standard deviation.

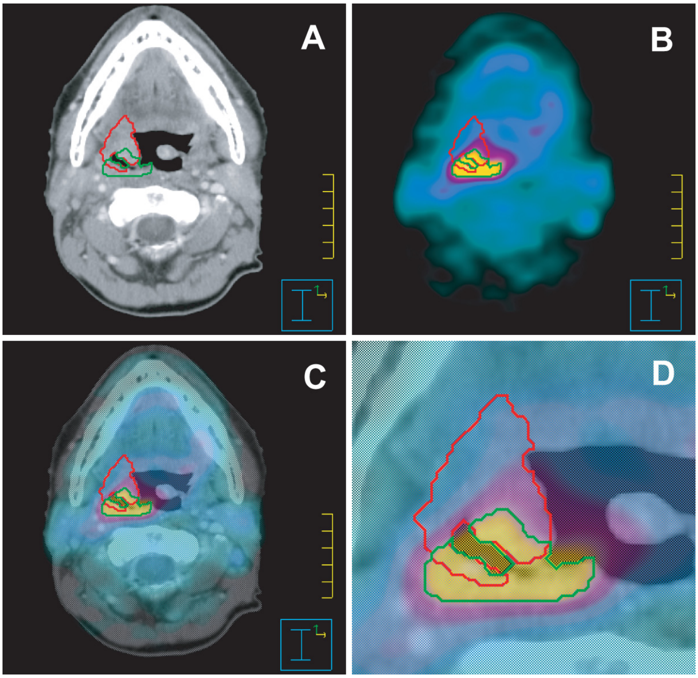
The mean absolute volumes obtained with the various delineation procedures are shown in table 7.2 and figure 7.2. GTV_{CT} and GTV_{VIS} yielded comparable volumes but the three threshold-based PET-GTVs (40%, 50% and SBR) all were smaller than the CT-based GTV ($p \leq 0.0001$ for all comparisons). Furthermore, $GTV_{50\%} < GTV_{40\%} < GTV_{VIS}$ ($p \leq 0.0003$), indicating that these methods resulted in significantly differently sized GTVs. The mean volumes of $GTV_{50\%}$ and GTV_{SBR} were very similar.

The results of the overlap analyses are also shown in table 7.2. The mean OF_{CT} varied from 0.39 to 0.61, depending on the segmentation tool used. The mean OF_{PET} varied from 0.66 to 0.85. A clear trend was observed with OF_{CT} decreasing and OF_{PET} increasing, from GTV_{VIS} to $GTV_{40\%}$ to $GTV_{50\%}$, GTV_{SBR}. This indicates that in this order the PET volume not only decreased but was

Additional tumor detected with FDG-PET				
Cases where >10% of GTV_{PET} is located outside the GTV_{CT}-volume				
	GTV_{VIS}	**$GTV_{40\%}$**	**$GTV_{50\%}$**	**GTV_{UCL}**
Oral cavity/oropharynx	30 (83%)	23 (68%)	17 (50%)	15 (43%)
Larynx/hypopharynx	35 (85%)	29 (74%)	20 (51%)	22 (55%)
Total	65 (84%)	52 (71%)	37 (51%)	37 (49%)
Cases where >20% of GTV_{PET} is located outside the GTV_{CT}-volume				
	GTV_{VIS}	**$GTV_{40\%}$**	**$GTV_{50\%}$**	**GTV_{UCL}**
Oral cavity/oropharynx	19 (53%)	19 (56%)	8 (24%)	9 (26%)
Larynx/hypopharynx	30 (73%)	24 (62%)	13 (33%)	14 (35%)
Total	49 (64%)	43 (59%)	21 (29%)	23 (31%)

Table 7.3
The number of cases where FDG-PET detected a significant additional tumor volume (GTV), that was not detected by CT.

Figure 7.3
CT scan (A), corresponding FDG-PET scan (B), fused image (C) and enhanced detail of fused image (D) of a patient with a T3N2cM0 oropharyngeal carcinoma, showing tumor tissue delineated by FDG-PET, but not by CT. Indicated are GTV_{CT} *(red) and* GTV_{SBR} *(dark green).*



also increasingly incorporated within the CT volume. All overlap fractions were significantly different from each other, except overlap fractions of $GTV_{50\%}$ versus GTV_{SBR}. The mean GTV fraction delineated by CT but not by PET, $1 \cap OF_{CT}$, varied from 0.39 to 0.61. The mean GTV fraction delineated by PET but not by CT, $1 - OF_{PET}$, varied from 0.15 to 0.34. The latter is further detailed in table 7.3, categorized for tumor subsite. The segmentation tools performed very differently with no clear difference by tumor site or stage and all methods resulted in a large percentage of patients having more than 20% of the GTV_{PET} outside the GTV_{CT}-domain. $GTV_{50\%}$ and GTV_{SBR} resulted in less PET-volume outside the CT-volume than GTV_{VIS} and $GTV_{40\%}$. An example of tumor tissue delineated by FDG-PET but not on CT is shown in figure 7.3.

The absolute volumes OF_{CT} and OF_{PET} of the methods $GTV_{50\%}$ and GTV_{SBR} were similar. The overlap fraction of $GTV_{50\%}$ relative to GTV_{CT}, versus the overlap fraction of GTV_{SBR} relative to GTV_{CT} showed a strong correlation between the two methods (Pearson correlation r=0.85, $p<0.0001$). Nevertheless, the overlap of $GTV_{50\%}$ and GTV_{SBR} relative to each other showed a similarity less than 90% in 26 cases. This indicates that, although the $GTV_{50\%}$ and GTV_{SBR} segmentation tools yield similar average GTV volumes and overlap fractions, on an individual patient basis, there is a geographical mismatch in a substantial number of cases.

7.4 Discussion

In this study we compared five segmentation tools for FDG-PET based target volume definition in a large cohort of head and neck cancer patients. There were three important observations. First, the GTV_{SUV} method using a fixed threshold of 2.5 failed to provide successful delineation in a large number of cases. Second, the volume and shape of the GTV on PET largely depends on the segmentation tool used. Third, PET frequently detected extension of tumor tissue outside the GTV_{CT}, regardless of the applied segmentation method.

Segmentation using a SUV-threshold of 2.5 resulted in an unsatisfactory large GTV in nearly half of the patients. In the remaining patients it resulted in a volume that was larger than the corresponding CT-volume, whilst the other automated segmentation tools produced volumes smaller than GTV_{CT}. The SUV of 2.5 was chosen as an arbitrary cut-off value as it represents the level that some reports use to consider a lesion as malignant, e.g. when staging non-small cell lung cancer (14,15), although the use of any SUV to differentiate benign from malignant is highly questionable (26). Furthermore, extrapolation to tumor sites other than lung cancer is even more debatable. Lung tumors are frequently much larger than head and neck tumors, and lung tumors are often surrounded by large areas of very low FDG activity (i.e. normal lung tissue) whereas in head and neck cancer there is always a significant amount of background activity, because of physiological FDG-uptake in surrounding muscular tissue. For segmentation of primary lung cancer, Nestle et al. have found the value of 2.5 for SUV satisfactory (10), but they also reported that when attempting to delineate a lesion surrounded by tissue with significant background activity that the GTV_{SUV} tool was not suitable. Based on our results, we conclude that this method is not useful for automated target definition of head and neck cancer.

The other PET segmentation tools were successful in most, or all (GTV_{VIS}), cases. We found no explanation why a $GTV_{40\%}$ and $GTV_{50\%}$ could not be generated in four cases. Miller et al. reported that thresholding at 40% only works when the SBR is larger than 10 (20). These four patients all had a SBR below 10 (range 3.6 to 6.4), However, 31 other patients who did have a satisfactory $GTV_{40\%}$ and $GTV_{50\%}$ also had a SBR below 10. We can not conclude that it was a low SBR that caused the unsatisfactory GTVs.

The mean absolute GTV volume derived from visual interpretation of PET-images was comparable to GTV_{CT}. All threshold-based PET segmentation tools, besides the rejected SUV-based segmentation, resulted in volumes smaller than GTV_{CT}. This was also described by Daisne et al. (27), who compared the role of co-registered CT, MRI and FDG-PET in delineating the primary tumor in laryngeal cancer patients prior to laryngectomy, with contouring based on the GTV_{SBR} tool. Compared to the reference surgical specimen all modalities overestimated the tumor extension, but GTV_{SBR} came closest at depicting the true tumor volume. The differences were significant, with an average GTV on histological examination of 12.6 cm^3, whereas averages for PET, CT and MRI were 16.3 cm^3, 20.8 cm^3 and 23.8 cm^3, respectively. We also observed this effect for CT and GTV_{SBR} in our patients with cancer of the larynx, oral cavity, oropharynx and hypopharynx.

Most other studies comparing GTV definition using FDG-PET against CT or MRI in head and neck cancer have used visual interpretation of PET images (5,9,28). All studies demonstrated significant differences between PET and CT volumes in a large proportion of the patients. Compared to the GTV_{CT}, the average GTV_{PET} was smaller in one study (5), of similar size in one study (9) and in another study larger in 40% of the cases, but smaller in the remaining 60% (28). These variable results may, at least partly, reflect the subjective nature and operator dependency of the visual interpretation method. Paulino et al. used a fixed threshold of 50% of the maximum signal intensity (21). The resulting GTV_{PET} was almost a factor two smaller relative to the GTV_{CT} which is in good agreement with our results.

Despite the fact that on average threshold-based GTV_{PET} volumes were smaller than GTV_{CT}, in many cases a significant part of the GTV_{PET} was located outside the volume defined as GTV_{CT}. This suggests that GTV definition using PET may include tumor extension that are not unequivocally depicted on CT. Note that our GTV_{CT} already included tumor extensions found at clinical examination and examination under anesthesia. As FDG may also accumulate in inflammatory tissue, this is a potential caveat resulting in larger PET-based GTV not due to cancer. The relevance of peritumoral inflammation needs further evaluation, as this could not be discriminated in the current study.

FDG-PET may also underestimate tumor volume in specific situations. Tumor parts that are small, or that are located in a background with a relatively intense signal, could be missed. Both Daisne et al. and Ng et al. observed that, like CT and MRI, FDG-PET failed to identify superficial mucosal tumor extension (27, 29). Nevertheless, using PET, Ng et al. missed only one small tumor location in a series of 124.

Given the different characteristics of GTV delineation with CT and PET, PET-based target volume definition can be used as complementary information to the conventional methods, but at present should not replace the CT-based volume until the smaller PET-based GTVs have been proven to be oncologically safe. This proof should ideally be acquired through histological validation studies, for then one can reliably decide how to use the additional information. GTV may be reduced using PET when dubious densities as seen on CT prove to be FDG-negative.

This will reduce overtreatment and possibly reduce side-effects. GTV may be expanded based on PET, when additional FDG-positive locations are not explained by clinically evident benign inflammation. This will reduce the risk of geographical misses. We envisage that the future role of PET in target volume definition will be as a complementary tool, adding to other imaging and clinical information.

The PET delineation methods resulted in GTV volumes with significant differences. Only the $GTV_{50\%}$ and GTV_{SBR} methods produced comparable volumes and overlap fractions. These segmentation tools, however, were not equivalent as geographical similarity between $GTV_{50\%}$ and GTV_{SBR} proved to be less than 90% in 26 of 73 cases. This might be explained by the thresholds that were generated by the GTV_{SBR} algorithm, which ranged from 45.8% to 63.7%. This illustrates that for an individual patient $GTV_{50\%}$ and GTV_{SBR} were not interchangeable. The differences between the other PET-based methods were larger both with regard to volume and overlap. This underlines the need for validated GTV definition with PET.

All segmentation tools have inherent limitations. The main weaknesses of GTV_{VIS} are that the resulting GTV is strongly influenced by the window-level setting of the data set and that it is a pure subjective approach, leading to substantial intra- and inter-observer variability (30). Using a fixed threshold (i.e. $GTV_{40\%}$ or $GTV_{50\%}$) as advocated by many research groups (15-21) is debatable. This seems to perform reasonably well in phantom-based experiments using symmetrical volumes with homogeneous activity and a sharp demarcation from the background activity. However, tumors may display a heterogeneous distribution of radioactivity. They may also be more complex in shape, which may limit the performance of the segmentation tool in the clinical setting (31). Furthermore, these methods imply that there is no dependency on the background activity. When trying to apply a fixed threshold on Daisne's laryngectomy data, thresholds ranging from 36% to 73% were necessary to fit the true tumor volume on histology (30,32). The GTV_{SBR} tool needs to be calibrated to the specific institutional image acquisition and reconstruction settings and it is not ideal for low SBR images (33). However, it is the only tool that has been the subject of a histological validation study, and its threshold is adapted to the signal to background ratio of an individual patient. A novel iterative method for lesion delineation and volumetric quantification has recently been presented, whereby the background is subtracted from the signal, which makes it independent of the signal-to-background ratio, and seems to be a more robust segmentation tool (34). It will be of great value to investigate its performance in head and neck cancer patients. Alltogether, a method that is not dependent on observer variations and SBR ratio, and that has been validated properly, is preferable. These criteria apply best to the GTV_{SBR} segmentation tool.

7.5 Conclusions

This study shows that FDG-PET may have important consequences for GTV definition, but that the choice of a segmentation tool for target volume definition of head and neck cancer based on PET images is not trivial. The absolute PET volume is dependent on the segmentation method used. Delineation using an SUV value of 2.5 is insufficient, and the other evaluated methods show inconsistencies. The SBR method seems preferable, since it uses a threshold adapted to the signal to background ratio of an individual patient and it does not depend on observer variability.

In general, PET volumes were smaller than CT volumes, but PET also identified possible tumor areas that were not contoured by the conventional CT-based method. This could potentially improve the accuracy of GTV definition. Additional histological validation studies are necessary before routine usage of FDG-PET data to optimize CT-derived target volumes.

7.6 References

1. Rischin D, Hicks RJ, Fisher R, et al. Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: a substudy of Trans-Tasman Radiation Oncology Group Study 98.02. J Clin Oncol 2006;24:2098-2104.
2. Ling CC, Humm J, Larson S, et al. Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. Int J Radiat Oncol Biol Phys 2000;47:551-560.
3. van Baardwijk A, Baumert BG, Bosmans G, et al. The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning. Cancer Treat Rev 2006;32:245-260.
4. Bradley JD, Perez CA, Dehdashti F, et al. Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. J Nucl Med 2004;45 Suppl 1:96S-101S.
5. Heron DE, Andrade RS, Flickinger J, et al. Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report. Int J Radiat Oncol Biol Phys 2004;60:1419-1424.
6. Hebert ME, Lowe VJ, Hoffman JM, et al. Positron emission tomography in the pretreatment evaluation and follow-up of non-small cell lung cancer patients treated with radiotherapy: preliminary findings. Am J Clin Oncol 1996;19:416-421.
7. Kiffer JD, Berlangieri SU, Scott AM, et al. The contribution of 18F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer. Lung Cancer 1998;19:167-177.

8. Nestle U, Walter K, Schmidt S, et al. 18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. Int J Radiat Oncol Biol Phys 1999;44:593-597.

9. Nishioka T, Shiga T, Shirato H, et al. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. Int J Radiat Oncol Biol Phys 2002;15:1051-1057.

10. Nestle U, Kremp S, Schaeffer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J Nucl Med 2005;46:1342-1348.

11. Messa C, Ceresoli GL, Rizzo G, et al. Feasibility of [18F]FDG-PET and coregistered CT on clinical target volume definition of advanced non-small cell lung cancer. Q J Nucl Med Mol Imaging 2005;49:259-266.

12. Riegel AC, Berson AM, Destian S, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. Int J Radiat Oncol Biol Phys 2006;65:726-732.

13. Hong R, Halama J, Bova D, et al. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. Int J Radiat Oncol Biol Phys 2007;67:720-726.

14. Paulino AC, Johnstone PA. FDG-PET in radiotherapy treatment planning: Pandora's box? Int J Radiat Oncol Biol Phys 2004;59:4-5.

15. Bradley J, Thorstad WL, Mutic S, et al. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2004;59:78-86.

16. Brianzoni E, Rossi G, Ancidei S, et al. Radiotherapy planning: PET/CT scanner performances in the definition of gross tumour volume and clinical target volume. Eur J Nucl Med Mol Imaging 2005;32:1392-1399.

17. Ciernik IF, Dizendorf E, Baumert BG, et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. Int J Radiat Oncol Biol Phys 2003;57:853-863.

18. Giraud P, Grahek D, Montravers F, et al. CT and (18)F-deoxyglucose (FDG) image fusion for optimization of conformal radiotherapy of lung cancers. Int J Radiat Oncol Biol Phys 2001;49:1249-1257.

19. Mah K, Caldwell CB, Ung YC, et al. The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. Int J Radiat Oncol Biol Phys 2002;52:339-350.

20. Miller TR, Grigsby PW. Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. Int J Radiat Oncol Biol Phys 2002;53:353-359.

21. Paulino AC, Koshy M, Howell R, et al. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2005;61:1385-1392.

22. Daisne JF, Sibomana M, Bol A, et al. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother Oncol 2003;69:247-250.

23. Vogel WV, Wensing BM, van Dalen JA, et al. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. Eur J Nucl Med Mol Imaging 2005;32:1276-1282.

24. Vogel WV, Schinagl DA, van Dalen JA, et al. Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. Q J Nucl Med Mol Imaging 2007 (in press)

25. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. J Nucl Med 2004;45:1431-1434.

26. Westerterp M, Pruim J, Oyen W, et al. Quantification of FDG PET studies using standardized uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. Eur J Nucl Med Mol Imaging 2007;34:392-404.

27. Daisne JF, Duprez T, Weynand B, et al. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 2004;233:93-100.

28. Scarfone C, Lavely WC, Cmelak AJ, et al. Prospective feasibility trial of radiotherapy target definition for head and neck cancer using 3-dimensional PET and CT imaging. J Nucl Med 2004;45:543-552.

29. Ng SH, Yen TC, Liao CT, et al. 18F-FDG PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of 124 patients with histologic correlation. J Nucl Med 2005;46:1136-1143.

30. Grégoire V, Bol A, Geets X, et al. Is PET-based treatment planning the new standard in modern radiotherapy? The head and neck paradigm. Semin Radiat Oncol 2006;16:232-238.

31. Hoffmann AL, van Dalen J, Lee J, et al. Assessment of 18F PET signals for automatic target volume definition in radiotherapy treatment planning: regarding Davis et al. Radiother Oncol 2007;83:102-103.

32. Grégoire V, Daisne JF, Geets X. Comparison of CT- and FDG-PET-defined GT: in regard to Paulino et al. Int J Radiat Oncol Biol Phys 2005;63:308-309.

33. Grégoire V, Haustermans K, Geets X, et al. PET-based treatment planning in radiotherapy: a new standard? J Nucl Med 2007;48 Suppl 1:68S-77S.

34. van Dalen JA, Hoffmann AL, Dicken V, et al. A novel iterative method for lesion delineation and volumetric quantification with FDG-PET. Nucl Med Comm 2007 (in press)

8

FLT-PET does not discriminate between reactive and metastatic lymph nodes in primary head and neck cancer patients

Esther G.C. Troost [1]
Wouter V. Vogel [2]
Matthias A.W. Merkx [3]
Piet J. Slootweg [5]
Henri A.M. Marres [4]
Wenny J.M. Peeters [1]
Johan Bussink [1]
Albert J. van der Kogel [1]
Wim J.G. Oyen [2]
Johannes H.A.M. Kaanders [1]

Departments of Radiation oncology [1], Nuclear medicine [2], Oral and maxillofacial surgery [3], Otorhinolaryngology, head and neck surgery [4] and Pathology [5], Radboud University Nijmegen Medical Centre, the Netherlands

Journal of Nuclear Medicine, 2007
Volume 48 (5), pages 726-735

Abstract

Repopulation of clonogenic tumor cells is inversely correlated with radiation treatment outcome in head and neck squamous cell carcinomas. A functional imaging tool to assess the proliferative activity of tumors could improve patient selection for treatment modifications and could be used for early treatment response evaluation. The PET tracer ^{18}F-fluor-deoxy-thymidine (FLT) can image tumor cell proliferation prior to and during radiotherapy, and it may provide biological tumor information useful in radiotherapy planning. In the present study, the value of FLT-PET in determining the lymph node status in squamous cell carcinoma of the head and neck was assessed, with pathology as the gold standard.

Methods: Ten patients with newly diagnosed stage II-IV squamous cell carcinoma of the head and neck underwent FLT-PET prior to surgical tumor resection with lymph node dissection. Emission FLT-PET and CT images of the head and neck were recorded and fused, and standardized uptake values (SUV) were calculated. From all 18 FLT-PET positive lymph node levels and from eight FLT-PET negative controls, paraffin embedded lymph node sections were stained and analyzed for the endogenous proliferation marker Ki-67 and for the pre-operatively administered proliferation marker iododeoxyuridine. Sensitivity, specificity, positive predictive value and negative predictive value were calculated for FLT-PET.

Results: Primary tumor sites were the oral cavity (7), larynx (2) and maxillary sinus (1). Nine of the ten patients examined had FLT-PET positive lymph nodes (SUV_{mean}: median 1.2, range 0.8-2.9), but only three of these patients had histologically proven metastases. All metastatic lymph nodes showed Ki-67 and iododeoxyuridine staining in tumor cells. In the remaining seven patients, there was abundant Ki-67 and iododeoxyuridine staining of B-lymphocytes in germinal centers in PET positive lymph nodes explaining the high rate of false positive findings. Sensitivity, specificity, positive and negative predictive values of FLT-PET were 100%, 16.7%, 37.5% and 100%, respectively.

Conclusion: In head and neck cancer patients, FLT-PET showed uptake in metastatic as well as in non-metastatic reactive lymph nodes, the latter due to reactive B-lymphocyte proliferation. Because of the low specificity, FLT-PET is not suitable for assessment of pre-treatment lymph node status. This observation may also negatively influence the utility of FLT-PET for early treatment response evaluation of small metastatic nodes.

8.1 Introduction

Lymph node involvement in squamous cell carcinoma of the head and neck is a poor prognostic indicator, reducing cure rate by almost 50% (1). The standard diagnostic work-up for assessing cervical lymph node status is performed by computed tomography (CT) or magnetic resonance imaging (MRI). The sensitivity (50-80%) and specificity (70-90%) of CT and MRI are relatively limited, and are comparable (2,3). For marginally enlarged lymph nodes, examination by ultrasound imaging (US) with fine needle aspiration cytology is superior to CT and MRI if performed by an experienced radiologist (sensitivity and specificity up to 76% and 100%, respectively) (4,5). More recently, a number of studies have been performed to assess the value of positron emission tomography with ^{18}F-fluor-deoxy-glucose (FDG-PET) for cervical lymph node staging (6-11). The results of these studies indicate that the performance of FDG-PET is not clearly superior to US, CT or MRI. Therefore, FDG-PET is generally not considered as part of the standard work-up for head and neck cancer patients for this indication.

An additional biologic factor of prognostic relevance is tumor cell proliferation. Head and neck squamous cell carcinomas may show accelerated repopulation of clonogenic tumor cells during the course of radiation therapy, and this is related to poor treatment outcome (12-15). Several treatment modifications have been developed to counteract this phenomenon, such as accelerated radiotherapy and inhibition of the epidermal growth factor receptor (EGFR), but at the cost of increased toxicity for the patient (16-18). Hence, careful patient selection for these treatment strategies is required ensuring maximal patient benefit, while preventing undue toxicity and costs. A diagnostic tool to identify lymph node metastases with high accuracy that can also provide information on the proliferative activity of the tumor could be of great value for treatment selection and radiotherapy planning.

Shields et al. introduced the novel PET-tracer ^{18}F-fluor-deoxy-thymidine (FLT) that is monophosphorylated by the cytosolic enzyme thymidine kinase 1 (TK1) and trapped intracellularly (19). TK1 activity is increased during DNA synthesis, and FLT trapping is related to TK1 activity and thus to proliferation. A number of studies evaluated the usefulness of FLT in assessing tumor cell proliferation in the primary tumor, most of them including a comparison with FDG-PET (20-31). Several studies validated FLT tracer uptake with the proliferation marker Ki-67 in primary tumor resection material or biopsies (23,30-35). A single study on laryngeal carcinoma by Cobben et al. compared FLT- with FDG-PET for imaging of the primary tumor without histological verification (27). Only three studies, two on breast carcinoma and one on thoracic tumors, have validated FLT-PET versus histopathology for the detection of metastatic lymph nodes (20,30,31).

Different markers have been used for histological assessment of proliferation. These include endogenous markers such as Ki-67, proliferating nuclear antigen (PCNA) and members of the cyclin group, or intravenous administration of the thymidine analogues bromodeoxyuridine (BrdUrd) and iododeoxyuridine (IdUrd) (36-38). The latter have a short half-life and are rapidly incorporated in the DNA of S-phase cells (39). For immunohistochemical validation of FLT these thymidine analogues seem most suitable because TK1 activity is increased mainly during DNA synthesis.

Thus far, validation of the PET-tracer FLT has mainly focused on primary tumor sites and only recently this was expanded to determining lymph node status. Characterization of both the primary tumor and the lymph nodes is compulsory for selection of the treatment strategy in patients with squamous cell carcinomas of the head and neck. The aim of this study was to determine the value of FLT-PET for assessment of the cervical lymph node status and proliferative activity with histological evaluation as the gold standard.

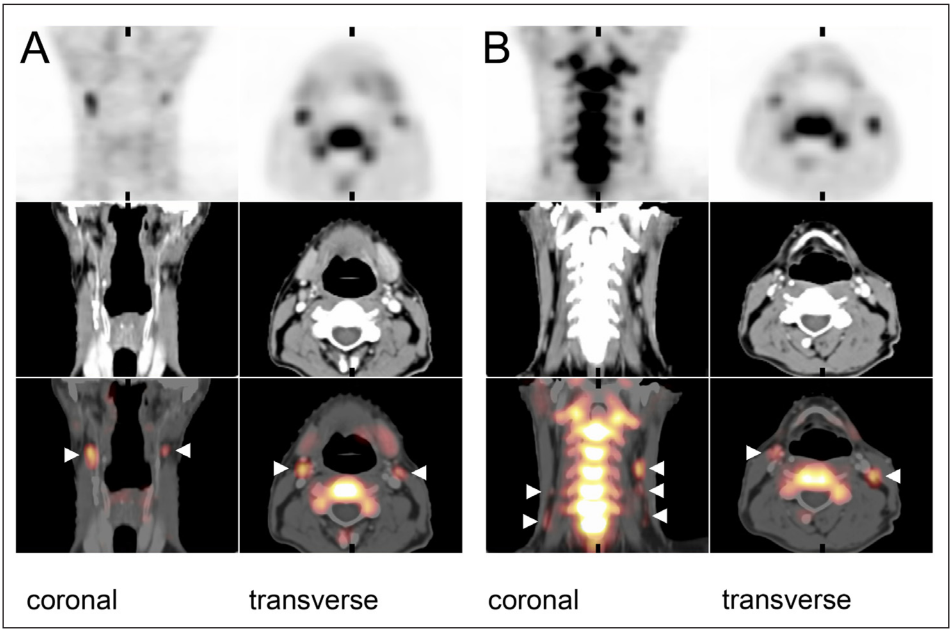


Figure 8.1
FLT-PET-CT images of patient number 9 (pT2N0M0 oral cavity carcinoma). The upper panels show the PET-images, the middle panels the CT-images and the lower panels the fusion of both image modalities. Cervical lymph nodes with increased FLT uptake are found bilaterally in level II (A: indicated by white arrows) and in levels III and IV (B: white arrows). All lymph nodes detected with FLT in this example were false-positive for metastasis, due to uptake in proliferating B-lymphocytes in reactive germinal centers.

8.2 Materials and methods

Patients

Ten patients with newly diagnosed stage II-IV primary squamous cell carcinoma of the head and neck, awaiting surgical tumor and lymph node resection, were included in this study after giving written informed consent. The study was approved by the Institutional Review Board of the Radboud University Nijmegen Medical Centre, the Netherlands.

FLT synthesis

FLT was obtained commercially from the Cyclotron B.V., VU Medical Centre, Amsterdam, The Netherlands. Synthesis was performed according to the method of Machulla et al. (40). In brief, FLT was produced by ^{18}F-fluorination of the 4,4′-dimethoxytrityl protected anhydrothymidine, followed by a deprotection step. After purification by reversed phase high performance liquid chromatography, the product was made isotonic and passed through a 0.22 µm filter. FLT was produced with a radiochemical purity of >95% and specific activity of >10 TBq/mmol.

PET/CT acquisition

Prior to surgical tumor resection, integrated PET and CT images were acquired with either a hybrid PET/CT, or with software fusion of dedicated PET and CT images. All scans were performed with the patient positioned in a rigid customized mask covering the head/neck area in order to increase position accuracy and to reduce movement artefacts during PET-scanning. Hybrid PET/CT images were acquired using a Siemens Biograph Duo scanner (Siemens/CTI, Knoxville, Tennessee, USA). Emission images of the head and neck area were recorded 60 minutes after intravenous injection of 250 MBq FLT, with 7 minutes per bed position in 3D mode. PET images were reconstructed using the OSEM iterative algorithm with parameters optimized for the head and neck area (i.e. 4 iterations, 16 subsets and 5 mm 3D Gaussian filter), with correction for photon attenuation. In addition, CT images were acquired with 120 mAs, 130 kV, and slice width 3 mm, with i.v. contrast in the venous phase, for anatomical correlation and attenuation correction purposes.

Dedicated PET images were acquired using a Siemens ECAT Exact scanner (Siemens/CTI, Knoxville, Tennessee, USA). Emission and transmission images of the head and neck area were recorded 60 minutes after intravenous injection of 250 MBq FLT, with 5 minutes per bed position in 3D mode for emission and 3 minutes per bed position in 2D mode for transmission. PET image reconstruction was identical to images from hybrid PET/CT. Dedicated CT images were acquired using a Philips AcQsim CT scanner (Philips, Cleveland, USA), with the same acquisition parameters as CT images from hybrid PET/CT. PET and CT image sets were anatomically co-registered using iterative closest point based optimization of surface maps derived from PET transmission and CT images, with an average registration accuracy of 3 mm (41).

PET analysis

Before analyzing the histological sections, the combined PET/CT image sets were reviewed in consensus by two experienced observers blinded for all patient data. Images were scored for presence or absence of FLT-PET uptake. Lymph node levels were determined as described by Gregoire et al. (42). Maximum and mean standardized uptake values (SUV_{max} and SUV_{mean}) were calculated for visible lymph nodes. SUV_{mean} was calculated after constructing a region of interest at the 50% isocontour of the SUV_{max}.

Surgery

Twenty minutes before the start of surgery, 200 mg IdUrd (Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland), diluted in 100 ml NaCl 0.9%, was administered intravenously as a bolus injection. After resection, the neck dissection specimens were presented on a uniform left or right sided plate resembling the neck levels I-VI.

Immunohistochemical staining of IdUrd and Ki-67

From the ten patients, a total of 236 lymph nodes without metastases and 14 lymph nodes with metastatic involvement localized in 44 lymph node levels were removed. From these, paraffin blocks containing lymph nodes from 26 different lymph node levels were collected for this study. These included 18 FLT-PET positive lymph node levels and 8 randomly chosen levels that were FLT-PET negative.

From these blocks, five µm sections were cut and consecutive sections were stained for Ki-67 and IdUrd. Between all consecutive steps of the staining procedure, lymph node sections were rinsed with 0.1 M phosphate-buffered saline (Klinipath, Duiven, The Netherlands), pH 7.4. The staining procedures were performed at room temperature unless stated differently. The sections were deparaffinized and re-hydrated in histosafe and graded alcohols. For antigen retrieval, slides were heated (90 °C) in 10 mM citrate buffer pH 6.0 for 30 min. For the Ki-67 staining, sections were incubated with normal donkey serum 5% diluted in primary antibody diluent (PAD, Abcam, Cambridge, UK) at 37 °C for 30 min. Overnight, sections were incubated with mouse-anti-human-Ki-67 (Zymed Laboratories, San Francisco, CA), undiluted, at 4 °C.

For the IdUrd staining, sections were incubated with 2 N HCl for 30 min followed by incubation, with 0.1 M Borax for 15 min and with normal donkey serum 5% diluted in PAD at 37 °C for 30 min. Then, sections were incubated with mouse-anti-IdUrd (Caltag Laboratories, Burlingame, CA), diluted 1:3000, for 60 min. In both staining procedures, peroxidase was blocked with 3% H_2O_2 in methanol for 10 min. Next, all sections were incubated with donkey-anti-mouse-biotin (Jackson Immuno Research Laboratories, West Grove, PA), diluted 1:400, for 60 min and with ABC-reagent (Vector Laboratories, Burlingame, CA) for 30 min. Then, sections were rinsed with deionized water before incubation with diaminobenzidine (Zymed Laboratories) for 15 min. Finally, after rinsing with tap water and staining with hematoxylin (Klinipath) for 30 sec, sections were dehydrated and captured in mounting medium (Klinipath).

Pathology evaluation and assessment of proliferation

The removed lymph nodes were routinely stained for hematoxylin and eosin (H&E) and assessed for metastatic involvement by a pathologist. Next, the clinical investigators and an experienced pathologist (PS) reviewed all lymph node sections stained for Ki-67 and IdUrd. Based on FLT-PET images and pathologic findings, lymph nodes were assigned to one of three groups: true positive (FLT positive lymph node with histologically proven metastasis), true

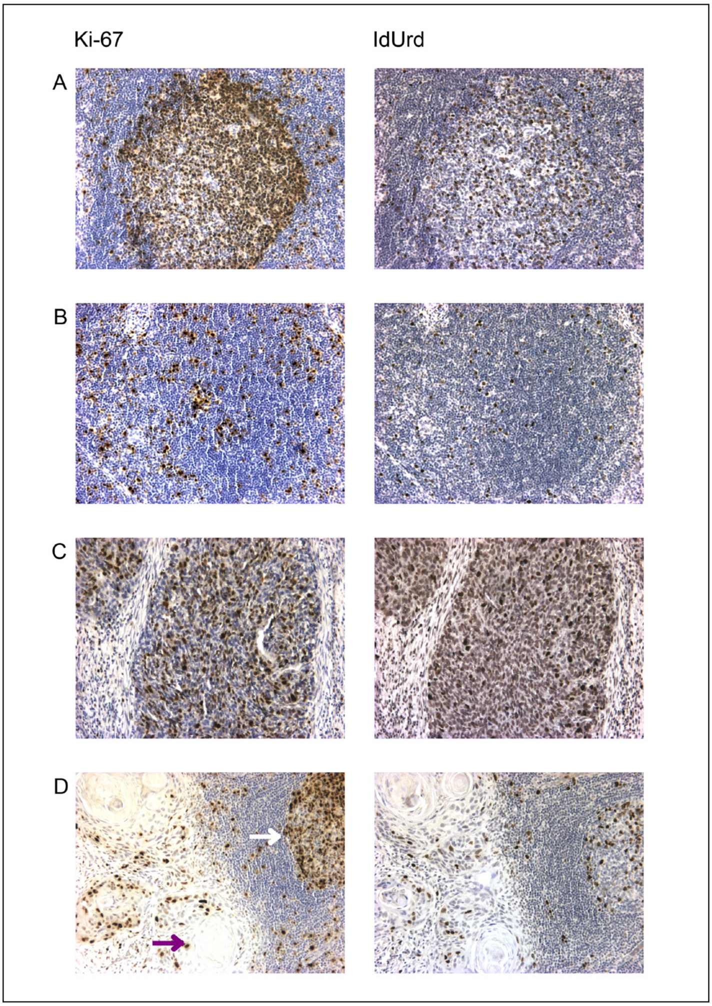


Figure 8.2
Ki-67 and IdUrd staining in: A: germinal center harboring proliferating B-lymphocytes and remaining lymphoid tissue; B: Remaining lymphoid tissue with proliferating lymphoid cells; C: metastasis of a squamous cell carcinoma of the maxillary sinus; D: micrometastasis with keratinization (indicated by purple arrow), fragment of a germinal center (white arrow) and surrounding lymphoid tissue.

negative (FLT negative lymph node without metastasis) and false positive (FLT positive lymph node without metastasis) or false negative (FLT negative lymph node with histologically proven metastasis).

Three histologically distinct areas were distinguished in the lymph node sections: germinal centers, metastatic tumor (if present) and remaining lymphoid tissue (figure 8.2). In these areas, Ki-67 and IdUrd positive and negative nuclei were counted using a grid with 25 fields placed in the eyepiece at 100x magnification (three randomly selected fields were analyzed in germinal centers, three fields in metastatic tissue and one field in remaining lymphoid tissue). The Ki-67 and IdUrd labeling index (LI) was determined as the number of positively stained nuclei relative to the total number of nuclei in a certain area.
In all lymph node sections, the total lymph node area and the (relative) area occupied by germinal centers and, if present, metastatic tumor were calculated. This was done by scanning the entire section under bright field microscopy and reconstructing a composite image of the complete lymph node using image analysis software (IPLab, Scanalytics Inc., Fairfax, VA). Masks were drawn on these scans indicating the total lymph node area, the germinal centers and metastatic tumor deposits. Next, using the image analysis software, the absolute and relative areas occupied by germinal centers and metastatic deposits were calculated.

SUV_{mean} versus Ki-67 and IdUrd staining

As a measure of total proliferative activity in the germinal centers of a lymph node section, the product of the Ki-67 or IdUrd LI and the absolute area occupied by the germinal centers was calculated. These parameters were called: $\text{Ki-67}_{\text{germinal center}}$ and $\text{IdUrd}_{\text{germinal center}}$. Similarly, as a measure of total proliferative activity in the entire lymph node, the sum of the products of LI and absolute area of germinal centers, remaining lymphoid tissue and, if present, metastatic deposits was calculated: $\text{Ki-67}_{\text{lymph node}}$ and $\text{Id}_{\text{Urdlymph}}$ node. These parameters were compared between true positive, false positive and true negative lymph nodes and were correlated with SUV_{mean}.

Statistical analysis

The ANOVA-test was used to assess differences in absolute area and Ki-67 and IdUrd LI between true negative, false positive and true positive lymph nodes. The t-test was applied for comparison of number and absolute area of germinal centers and for comparison of $\text{Ki-67}_{\text{germinal center}}$ and $\text{IdUrd}_{\text{germinal center}}$ between true negative and false positive lymph nodes. Correlations between $\text{Ki-67}_{\text{lymph node}}$ and $\text{IdUrd}_{\text{lymph node}}$ and SUV_{mean} were calculated using linear regression. All statistical analyses were calculated using GraphPad Prism (version 4.0a). A p-value ≤ 0.05 was regarded statistically significant.

Patients		US-FNA cytology	CT	MRI	Clinical stage	Path. stage	Procedure	Pathology LN*		LN+ on PET and available for path. assessm.		LN+ on PET but not available for path. assessm.
Nr	Site	(Side)						Left	Right	Left	Right	(Side)
1	Maxillary sinus	+ (left)	n.p.	pos.	T4N1M0	T4N1M0	TE[§] + MRND[‖] left	pos (1)	n.a.	1	n.a.	1 (right)
2	laryngeal	n.p.[†]	neg.	n.p.	T3N0M0	T3N0M0	TLE[¶] + bilateral NS[#]	neg	neg	1	1	2 (right)
3	laryngeal	n.p.	pos.	n.p.	T4N2cM0	T4N2cM0	TLE + bilateral RND[**]	pos (7)	pos (5)	2	2	1 (right)
4	Lower alveolar ridge	- (left)	neg.	n.p.	T4N2bM0	T4N0M0	TE + MRND right	n.a.	neg	n.a.	2	2 (left)
5	Floor of mouth	n.r.[‡]	n.p.	n.p.	T2N0M0	T2N0M0	TE + bilateral SND[††]	neg	neg	neg	neg	-
6	Tongue	- (right)	n.p.	n.p.	T3N0M0	T2N0M0	TE + SND right	n.a.	neg	-	5	-
7	Soft palate	n.p.	n.p.	n.p.	T2N0M0	T2N0M0	TE + SND left	neg	n.a.	neg	n.a.	1 (right)
8	Tongue	n.p.	n.p.	neg.	T3N0M0	T3N1M0	TE + SND right	n.a.	pos (1)	-	1	1 (right)
9	Tongue / floor of mouth	- (left)	n.p.	neg.	T4N0M0	T2N0M0	TE + SND left	neg	n.a.	3	n.a.	2 (right)
10	Tongue / floor of mouth	- (left)	n.p.	neg.	T2N0M0	T2N0M0	TE + SND left	neg	n.a.	2	-	-

Table 8.1
characteristics, diagnostic and therapeutic procedures and histopathology of lymph nodes
[*]LN = lymph node, [†]n.p. = not performed, [‡]n.r. = not representative, [§]TE = tumor excision, [‖]MRND = modified radical neck dissection, [¶]TLE = total laryngectomie, [#]NS = node sampling, [**]RND = radical neck dissection, [††]SND = selective neck dissection (level I-III).

8.3 Results

Patient characteristics are summarized in table 8.1. Four men and six women with a mean age of 59 years (range 43-80 years) were included. Primary tumor sites were oral cavity (7), larynx (2) and maxillary sinus (1). Preoperative staging of the neck was performed with ultrasound imaging (8 patients) – with fine needle aspiration cytology in 6 patients and without in two, with CT (3 patients) or MRI (4 patients). All patients underwent surgical tumor resection combined with assessment of cervical lymph node involvement according to the Dutch National Guidelines: six patients with clinically N0 oral cavity carcinoma underwent selective neck dissection of level I-III. In three patients a (modified) radical neck dissection was done because of pre-operatively proven cervical lymph node involvement. In one patient with a glottic laryngeal carcinoma without suspected lymph node involvement only sampling of level II and III nodes was performed. Neck surgery was performed unilaterally in seven patients and bilaterally in three patients.

PET imaging

FLT-PET scans were acquired consecutively on PET/CT (3) or separately on CT and PET (7). The median time interval between FLT-PET and surgery was 5 days (average 10 days, range 4-37 days). In all but one case, surgery was performed within 14 days after the PET-scan. In the case with the longest interval (37 days), the surgical resection of an oral cavity carcinoma was postponed because the patient underwent laser evaporation of a small laryngeal lesion first. In all but one patient (number 5), lymph nodes with increased mean SUV were detected (range 0.8-2.9, median 1.2, standard deviation (S.D.) 0.41), mostly in multiple lymph nodes. A typical FLT-PET-CT image is shown in figure 8.1.

Pathological evaluation

Routine pathology based on H&E staining revealed three patients to have metastatic cervical lymph node disease (numbers 1, 3, 8), whereby clinical examination and anatomical imaging had predicted this only in two patients (numbers 1 and 3). The third patient had a micrometastasis of less than 2 mm in one lymph node (number 8). Although preoperative staging of the neck revealed multiple ipsilateral enlarged lymph nodes in patient 4, final histology showed no signs of metastatic disease.

In total, paraffin-embedded sections containing 26 lymph node levels were selected for Ki-67 and IdUrd staining and analysis. Eighteen of these levels were positive on FLT-PET, eight levels were negative. As shown in table 8.2, comparison of FLT-PET results with pathology revealed six true positive, twelve false positive and eight true negative findings. There were no false negative FLT-PET studies. Based on these findings, sensitivity of FLT for determining lymph node status in head and neck cancer patients was 100%, specificity 40%, positive predictive

	Stage	**LN level**		**Pathology**		**Ki-67 and IdUrd staining**		
		FLT pos	FLT neg	Metastasis		Metastasis	Germinal centers	Remaining lymph node tissue
1	pT4N1M0	R[‡] II[||]		n.a.[†]				
			L III	-	(TN[¶])	-	+	+
		L[§] II		+	(TP[#])	+	-	+
2	pT3N0M0	R II		-	(FP[**])	-	++	+
		R III		n.a.				
		R IV		n.a.				
		L III		-	(FP)	-	++	+
3	pT4N2cM0	R II		n.a.				
		R III		+	(TP)	+	++	+
		R IV		+	(TP)	+	++	+
		L II		+	(TP)	+	-	-
		L IV		+	(TP)	+	+	+
4	pT4N0M0	R II		-	(FP)	-	++	+
			R III	-	(TN)	-	-	+
		R IV		-	(FP)	-	-	+
			R V	-	(TN)	-	++	+
		L II		n.a.				
		L IV		n.a.				
5	pT2N0M0		R I	-	(TN)	-	+	+
			L II	-	(TN)	-	+	++
6	pT2N0M0	R I		-	(FP)	-	++	++
		R II		-	(FP)	-	++	++
		R III		-	(FP)	-	+	+
7	pT2N0M0	R IV		n.a.				
			L I	-	(TN)	-	++	+
			L III	-	(TN)	-	++	+
8	pT3N1M0	R II		+	(TP)	+	++	+
		RIII		n.a.				
9	pT2N0M0	R II		n.a.				
		R IV		n.a.				
		L II		-	(FP)	-	++	+
		L III		-	(FP)	-	+	+
		L IV		-	(FP)	-	+	+
10	pT2N0M0		L I	-	(TN)	-	-	+
		L II		-	(FP)	-	+	++
		L III		-	(FP)	-	+	++

Table 8.2
Histological lymphnode assessment histopathology and Ki-67 and IdUrd staining
[*]LN = lymph node, [†]n.a. = not available, [‡]R = right, [§]L = left, [||]II = level indicated by Roman number, [¶]TN = true negative (FLT negative lymph node without metastasis), [#]TP = true positive (FLT positive l.n. with metastasis), [**]FP = false positive (FLT positive l.n. without metastasis)

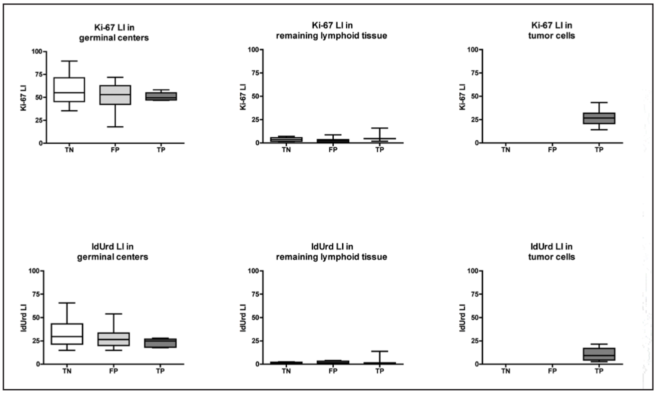


Figure 8.3
Results in germinal centers. Ki-67 and IdUrd LI in germinal centers, metastases and remaining lymphoid tissue for true negative (TN; FLT negative lymph node without metastasis), false positive (FP; FLT positive l.n. without metastasis) and true positive (TP; FLT positive l.n. with metastasis) lymph nodes.

value 33.3% and negative predictive value 100% on lymph node level. On patient level, after excluding patient 7 as no histological correlate for the FLT-positive lymph node was available, numbers were 100%, 16.7%, 37.5% and 100%, respectively.

Not all lymph nodes showing enhanced FLT-PET-uptake were removed during surgery. These lymph nodes were either not included in the neck dissection specimen (patient 3 and 8), or situated in the contralateral neck (patient 1, 2, 4, 7, 9). Based on the results of the standard diagnostic work-up and therapeutic guidelines, there was no indication for removal of these nodes.

Follow-up (median 13 months, range 11-15 months) has revealed a recurrent primary tumor in one patient treated with surgery and post-operative radiotherapy (pT2N0M0 carcinoma of the tongue). Until present, no lymph node recurrence has been observed in any of the patients.

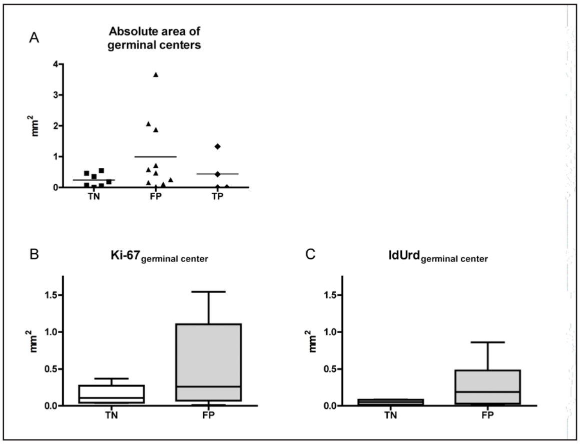


Figure 8.4
A: Absolute area (in mm^2) occupied by germinal centers in true negative (TN; FLT negative lymph node without metastasis), false positive (FP; FLT positive lymph node without metastasis) and true positive (TP; FLT positive lymph node with metastasis) lymph nodes. B and C: Ki-67$_{germinal\ center}$ and IdUrd $_{germinal\ center}$ in TN and FP lymph nodes as a measure of the total proliferative activity in germinal centers (calculated as area in mm^2).

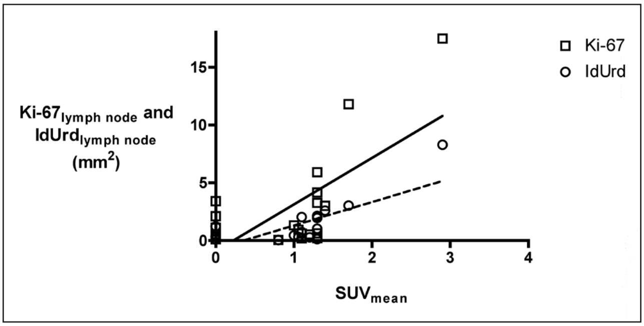


Figure 8.5
Scatterplot of Ki-67 lymph node and IdUrd lymph node as a measure of the total proliferative activity in the lymph node (calculated as area in mm^2*) versus SUV*$_{mean}$ *of FLT-PET. Solid line indicates linear best fit for Ki-67 and dashed line for IdUrd.*

Ki-67 and IdUrd staining

Ki-67 and IdUrd staining was present in metastatic tumor cells, germinal centers and in remaining lymphoid tissue as shown in figure 8.2. In almost all lymph nodes examined – both FLT-PET positive and negative – germinal centers staining positive for Ki-67 and IdUrd were present. Metastatic tumor cells in patients 1 and 3 had almost fully destroyed the lymph node architecture. In patient 8, the micrometastasis occupied only a small region of the affected lymph node leaving reactive germinal centers and remaining lymph node tissue unperturbed.

In germinal centers, metastases and remaining lymphoid tissue nuclei staining positive for Ki-67 and IdUrd were counted and the LI was calculated. Figure 8.3 shows the results for true negative, false positive and true positive lymph nodes. The median Ki-67 and IdUrd LI in the germinal centers was 53% and 26.9% with no difference between the three groups. In the remaining lymphoid tissue this was 3.6% and 1.6% respectively, also with no difference between the groups. In the metastases the median LI was 26.7% for Ki-67 and 9.3% for IdUrd. In all three areas and patient groups studied, Ki-67 LI was significantly higher compared to IdUrd LI ($p<0.0001$).

Germinal centers

The median number of germinal centers per lymph node was 9 (S.D. 6.7) in true negative lymph nodes, 20.5 (S.D. 27.0) in false positive nodes and 4 (S.D. 18.1) in true positive nodes. The difference in number of germinal centers between true negative and false positive lymph nodes was significant ($p=0.03$). Also the absolute area occupied by germinal centers was higher in false positive nodes relative to true negative nodes but the difference was only borderline significant ($p=0.06$), see figure 8.4.A. The total proliferative activity in the germinal centers expressed as Ki-67$_{\text{germinal center}}$ and IdUrd$_{\text{germinal center}}$ was higher in the false positive lymph nodes compared to the true negative nodes, although the difference did not reach statistical significance ($p=0.07$, respectively), see figures 8.4.B and 8.4.C.

Proliferative activity in lymph nodes and correlation with SUV_{mean}

The total proliferative activity in lymph nodes expressed as $Ki\text{-}67_{lymph\ node}$ and $IdUrd_{lymph\ node}$ was found to be significantly higher in the true positive lymph nodes as compared to false positive lymph nodes ($p=0.006$ and $p=0.05$, respectively). As shown in figure 8.5, there was a moderate but significant correlation between SUV_{mean} and $Ki\text{-}67_{lymph\ node}$ ($r^2=0.47$, $p=0.0009$) and between SUV_{mean} and $IdUrd_{lymph\ node}$ ($r^2=0.55$, $p=0.0004$).

8.4 Discussion

The PET tracer FLT for imaging of cell proliferation has been studied by various groups, correlating FLT with FDG (20-31). Buck et al. studied FLT-PET and FDG-PET in 47 patients with benign and malignant pulmonary nodules and found a high sensitivity of FLT for malignant primary tumors (sensitivity 90%), but not for mediastinal lymph node involvement (sensitivity 53%) or detection of lung metastasis (sensitivity 67%) (22). Twenty-one patients with primary or recurrent laryngeal carcinomas were studied by Cobben et al. Data on sensitivity and specificity for detection of the primary tumor were only reported for FDG and not for FLT. However, the standardized uptake value (SUV) for FLT was found to be significantly lower compared to the SUV for FDG and therefore the routine use of FLT-PET for detection of laryngeal carcinoma was not recommended (27). Additional studies correlated FLT with histological assessment of proliferation by Ki-67 labeling in fibrosarcoma, breast and lung cancer (23,30-35). In lung carcinoma, Buck et al. found that the Ki-67 LI correlated with FLT uptake (23). This finding has recently been confirmed by Yap et al. (31). In a fibrosarcoma xenograft, Leyton et al. found that FLT uptake and PCNA labelling index were linearly correlated (33). In contrast to these studies, Smyczek-Gargya et al. found no correlation between Ki-67 labelling index and FLT uptake in primary breast carcinoma (30). Although the results are not entirely consistent, these studies suggest that FLT-PET may be of value for the quantification of tumor cell proliferation.

In the current study, the value of FLT-PET in assessing the cervical lymph node status and proliferative activity of metastatic lymph nodes in squamous cell carcinoma of the head and neck was investigated. The endogenous proliferation marker Ki-67, which is expressed in the G1-, S-, G2- and M-phase of the cell cycle, was chosen for comparison with FLT-PET, because endogenous markers do not require intravenous administration and because recent studies validated FLT-PET with Ki-67 (23,30,31). In addition, the exogenous marker IdUrd was used. IdUrd is a robust and specific S-phase marker and it was hypothesized that this marker might correlate better with FLT uptake, because TK1 activity is increased mainly during DNA synthesis. In nine of the ten patients studied, increased FLT uptake was observed in the lymph nodes. Only three of these nine FLT-PET positive patients had metastatic nodal disease confirmed by histopathology, and two of them had already been detected by routine preoperative screening. In none of the FLT-PET negative lymph nodes metastatic disease was present.

As the number of false positive lymph nodes was high, further analysis was performed taking into account the architecture and proliferative state of the lymph nodes. In most lymph nodes evaluated intense staining of both Ki-67 and IdUrd was present in B-lymphocytes proliferating in germinal centers. Less intense staining was found in metastatic tumor cell deposits and the proliferative activity in remaining lymph node tissue was very low. In germinal centers, the labelling index of both markers was significantly higher compared to metastases or lymphoid tissue. There was no difference in the proliferative activity in germinal centers between true negative and false positive lymph nodes. However, false positive nodes on average contained a significantly greater number of germinal centers and these occupied a larger absolute area relative to the true negative nodes, although this latter finding was only borderline statistically significant. Also the product of Ki-67 LI and IdUrd LI and area occupied by germinal centers was higher in the false positive compared to the true negative lymph nodes. It is therefore likely that the active proliferation of B-lymphocytes in germinal centers is responsible for the false-positive FLT-PET results. This high proliferative activity of B-lymphocytes might also be responsible for FLT-PET positivity of the micrometastasis in patient 8.

Three other studies, two in breast cancer and one in thoracic tumors, compared FLT-PET with histopathology for the assessment of lymph node status (20,30,31). The study by Smyczek-Gargya et al. included 14 breast cancer patients, of whom eight had histologically proven axillary lymph node metastasis and seven were detected by FLT-PET (sensitivity 87.5% and specificity 100%) (30). In the study by Been et al. that included ten patients, only two of seven patients with histologically proven metastatic axillary lymph nodes were detected by FLT-PET (sensitivity 28.5% and specificity 100%) (20). Yap et al. studied 22 patients with thoracic tumors and reported sensitivity and specificity for detection of mediastinal lymph nodes by FLT to be 33.3% and 98.2%, respectively (31). The low sensitivity in some of these studies may be explained by the fact that some histological tumor types, such as mammary carcinoma, can exhibit limited proliferative activity. A second explanation might be that the metastatic tumor load of some of these lymph nodes was low.

All three studies reported a high specificity in contrast to the current study were specificity was only 16.7%. The localization of the lymph nodes may be of importance in explaining this discrepancy. Reactive lymph nodes in the head and neck area are found frequently as response to bacterial or viral infections, whereas reactive axillary lymph nodes are less common. Furthermore, patients with squamous cell carcinomas of the oral cavity – in this study the largest patient group – often present with non-healing ulcers accompanied with reactive lymph nodes. This is consistent with the observation that false positive lymph nodes on average contained a higher subvolume of germinal centers with very active proliferation of B-lymphocytes as compared to true negative nodes.

In this study, FLT-PET reached only low mean SUV in metastatic as well as in non-metastatic cervical lymph nodes compared to the SUV for FDG-PET generally reported in head and neck cancer (6-9). This is in agreement with the previous finding for primary laryngeal tumors by Cobben et al. (27). In accordance with studies discussed above, a significant correlation between SUV_{mean} and overall Ki-67 and IdUrd staining ($\text{Ki-67}_{\text{lymph node}}$ and $\text{IdUrd}_{\text{lymph node}}$) was found in this study (figure 8.5) (31,33). The correlation was strongest for IdUrd and we recommend this marker for future studies with FLT that include histological validation.

8.5 Conclusions

Although FLT-PET correctly identified all head and neck cancer patients with metastatic lymph nodes, the specificity and positive predictive value were low due to tracer-uptake in germinal centers of lymph nodes. Therefore, the use of FLT-PET for assessing pretreatment lymph node status and for determining the proliferative activity of affected lymph nodes is not encouraged in head and neck cancer patients. This also limits the usefulness of FLT-PET for radiation therapy planning and possibly for early treatment response evaluation of small metastatic lymph nodes. The value of FLT-PET for assessment of the proliferative state of the primary head and neck tumor and the relevance for radiotherapy planning is a topic of current investigations.

Acknowledgements

The authors would like to thank the nuclear medicine technologists for their assistance and excellent patient care. This work was supported by Junior Researcher Grant 2006-38 of the Radboud University Nijmegen Medical Centre and by EC FP6 funding (Biocare Contract no. LSHC-CT-2004-505785).

8.6 References

1. Brockstein B, Haraf DJ, Rademaker AW, et al. Patterns of failure, prognostic factors and survival in locoregionally advanced head and neck cancer treated with concomitant chemoradiotherapy: a 9-year, 337-patient, multi-institutional experience. Ann Oncol. 2004;15(8):1179-1186.
2. van den Brekel MW, Castelijns JA, Stel HV, Golding RP, Meyer CJ, Snow GB. Modern imaging techniques and ultrasound-guided aspiration cytology for the assessment of neck node metastases: a prospective comparative study. Eur Arch Otorhinolaryngol. 1993;250(1):11-17.
3. Wide JM, White DW, Woolgar JA, Brown JS, Vaughan ED, Lewis-Jones HG. Magnetic resonance imaging in the assessment of cervical nodal metastasis in oral squamous cell carcinoma. Clin Radiol. 1999;54(2):90-94.
4. Takes RP, Righi P, Meeuwis CA, et al. The value of ultrasound with ultrasound-guided fine-needle aspiration biopsy compared to computed tomography in the detection of regional metastases in the clinically negative neck. Int J Radiat Oncol Biol Phys. 1998;40(5):1027-1032.
5. van den Brekel MW, Castelijns JA, Stel HV, et al. Occult metastatic neck disease: detection with US and US-guided fine-needle aspiration cytology. Radiology. 1991;180(2):457-461.
6. Adams S, Baum RP, Stuckensen T, Bitter K, Hor G. Prospective comparison of 18F-FDG PET with conventional imaging modalities (CT, MRI, US) in lymph node staging of head and neck cancer. Eur J Nucl Med. 1998;25(9):1255-1260.
7. Dammann F, Horger M, Mueller-Berg M, et al. Rational diagnosis of squamous cell carcinoma of the head and neck region: comparative evaluation of CT, MRI, and 18FDG PET. AJR Am J Roentgenol. 2005;184(4):1326-1331.
8. Ng SH, Yen TC, Liao CT, et al. 18F-FDG PET and CT/MRI in oral cavity squamous cell carcinoma: a prospective study of 124 patients with histologic correlation. J Nucl Med. 2005;46(7):1136-1143.
9. Stuckensen T, Kovacs AF, Adams S, Baum RP. Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J Craniomaxillofac Surg. 2000;28(6):319-324.
10. Ng SH, Yen TC, Chang JT, et al. Prospective study of [18F]fluorodeoxyglucose positron emission tomography and computed tomography and magnetic resonance imaging in oral cavity squamous cell carcinoma with palpably negative neck. J Clin Oncol. 2006;24(27):4371-4376.
11. Wensing BM, Vogel WV, Marres HA, et al. FDG-PET in the clinically negative neck in oral squamous cell carcinoma. Laryngoscope. 2006;116(5):809-813.
12. Kim JJ, Tannock IF. Repopulation of cancer cells during therapy: an important cause of treatment failure. Nat Rev Cancer. 2005;5(7):516-525.

13. Petersen C, Zips D, Krause M, et al. Repopulation of FaDu human squamous cell carcinoma during fractionated radiotherapy correlates with reoxygenation. Int J Radiat Oncol Biol Phys. 2001;51(2):483-493.

14. Taylor JM, Withers HR, Mendenhall WM. Dose-time considerations of head and neck squamous cell carcinomas treated with irradiation. Radiother Oncol. 1990;17(2):95-102.

15. Withers HR, Taylor JM, Maciejewski B. The hazard of accelerated tumor clonogen repopulation during radiotherapy. Acta Oncol. 1988;27(2):131-146.

16. Kaanders JH, van der Kogel AJ, Ang KK. Altered fractionation: limited by mucosal reactions? Radiother Oncol. 1999;50(3):247-260.

17. Bonner JA, Harari PM, Giralt J, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. N Engl J Med. 2006;354(6):567-578.

18. Overgaard J, Hansen HS, Specht L, et al. Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHANCA 6 and 7 randomised controlled trial. Lancet. 2003;362(9388):933-940.

19. Shields AF, Grierson JR, Dohmen BM, et al. Imaging proliferation in vivo with [F-18]FLT and positron emission tomography. Nat Med. 1998;4(11):1334-1336.

20. Been LB, Elsinga PH, de Vries J, et al. Positron emission tomography in patients with breast cancer using (18)F-3’-deoxy-3’-fluoro-l-thymidine ((18)F-FLT)-a pilot study. Eur J Surg Oncol. 2006;32(1):39-43.

21. Buck AK, Halter G, Schirrmeister H, et al. Imaging proliferation in lung tumors with PET: 18F-FLT versus 18F-FDG. J Nucl Med. 2003;44(9):1426-1431.

22. Buck AK, Hetzel M, Schirrmeister H, et al. Clinical relevance of imaging proliferative activity in lung nodules. Eur J Nucl Med Mol Imaging. 2005;32(5):525-533.

23. Buck AK, Schirrmeister H, Hetzel M, et al. 3-deoxy-3-[(18)F]fluorothymidine-positron emission tomography for noninvasive assessment of proliferation in pulmonary nodules. Cancer Res. 2002;62(12):3331-3334.

24. Chen W, Cloughesy T, Kamdar N, et al. Imaging proliferation in brain tumors with 18F-FLT PET: comparison with 18F-FDG. J Nucl Med. 2005;46(6):945-952.

25. Choi SJ, Kim JS, Kim JH, et al. [18F]3’-deoxy-3’-fluorothymidine PET for the diagnosis and grading of brain tumors. Eur J Nucl Med Mol Imaging. 2005;32(6):653-659.

26. Cobben DC, Koopal S, Tiebosch AT, et al. New diagnostic techniques in staging in the surgical treatment of cutaneous malignant melanoma. Eur J Surg Oncol. 2002;28(7):692-700.

27. Cobben DC, van der Laan BF, Maas B, et al. 18F-FLT PET for visualization of laryngeal cancer: comparison with 18F-FDG PET. J Nucl Med. 2004;45(2):226-231.

28. Dittmann H, Dohmen BM, Paulsen F, et al. [18F]FLT PET for diagnosis and staging of thoracic tumours. Eur J Nucl Med Mol Imaging. 2003;30(10):1407-1412.

29. Francis DL, Freeman A, Visvikis D, et al. In vivo imaging of cellular proliferation in colorectal cancer using positron emission tomography. Gut. 2003;52(11):1602-1606.

30. Smyczek-Gargya B, Fersis N, Dittmann H, et al. PET with [18F]fluorothymidine for imaging of primary breast cancer: a pilot study. Eur J Nucl Med Mol Imaging. 2004;31(5):720-724.

31. Yap CS, Czernin J, Fishbein MC, et al. Evaluation of thoracic tumors with 18F-fluorothymidine and 18F-fluorodeoxyglucose-positron emission tomography. Chest. 2006;129(2):393-401.

32. Kenny LM, Vigushin DM, Al-Nahhas A, et al. Quantification of cellular proliferation in tumor and normal tissues of patients with breast cancer by [18F]fluorothymidine-positron emission tomography imaging: evaluation of analytical methods. Cancer Res. 2005;65(21):10104-10112.

33. Leyton J, Latigo JR, Perumal M, Dhaliwal H, He Q, Aboagye EO. Early detection of tumor response to chemotherapy by 3’-deoxy-3’-[18F]fluorothymidine positron emission tomography: the effect of cisplatin on a fibrosarcoma tumor model in vivo. Cancer Res. 2005;65(10):4202-4210.

34. Muzi M, Vesselle H, Grierson JR, et al. Kinetic analysis of 3’-deoxy-3’-fluorothymidine PET studies: validation studies in patients with lung cancer. J Nucl Med. 2005;46(2):274-282.

35. Vesselle H, Grierson J, Muzi M, et al. In vivo validation of 3’deoxy-3’-[(18)F]fluorothymidine ([(18)F]FLT) as a proliferation imaging tracer in humans: correlation of [(18)F]FLT uptake by positron emission tomography with Ki-67 immunohistochemistry and flow cytometry in human lung tumors. Clin Cancer Res. 2002;8(11):3315-3323.

36. Begg AC, Haustermans K, Hart AA, et al. The value of pretreatment cell kinetic parameters as predictors for radiotherapy outcome in head and neck cancer: a multicenter analysis. Radiother Oncol. 1999;50(1):13-23.

37. Brown DC, Gatter KC. Ki67 protein: the immaculate deception? Histopathology. 2002;40(1):2-11.

38. Kelman Z. PCNA: structure, functions and interactions. Oncogene. 1997;14(6):629-640.

39. Hughes WL, Commerford SL, Gitlin D, et al. Deoxyribonucleic Acid Metabolism In Vivo: I. Cell Proliferation And Death As Measured By Incorporation And Elimination Of Iododeoxyuridine. Fed Proc. 1964;23:640-648.

40. Machulla HJ. Simplified labeling approach for synthesizing 3’-deoxy-3’-18F-fluorothymidine. J Radioanal Nucl Chem. 2000;24:843-846.

41. Vogel WV, Schinagl DAX, van Dalen JA, Kaanders JHAM, Oyen WJG. Validated image fusion of dedicated PET and CT, for external beam radiation therapy in the head and neck area. Quarterly Journal of Nuclear Medicine. 2006; in press.

42. Gregoire V, Coche E, Cosnard G, Hamoir M, Reychler H. Selection and delineation of lymph node target volumes in head and neck conformal radiotherapy. Proposal for standardizing terminology and procedure based on the surgical experience. Radiother Oncol. 2000;56(2):135-150.

9

The role of PET/CT in the detection of recurrent colorectal cancer

Wouter V. Vogel [1]
Bas Wiering [2]
Frans H.M. Corstens [1]
Theo J.M. Ruers [2]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1] and Surgery [2],
Radboud University Nijmegen Medical Centre, the Netherlands

Cancer Imaging, 2005
Volume 5, pages S143-149.

Abstract

Many imaging modalities and scanning techniques, such as contrast enhanced CT, (dynamic) MRI and FDG-PET, are available for assessment of recurrent colorectal carcinoma. On top of this, integrated PET/CT is becoming increasingly available. Intuitively, a synergistic combination of scanning characteristics sounds promising.

However, the exact clinical value has not yet been fully established. The role of PET/CT image fusion must be weighed carefully against other available modalities. In this review we evaluate the potential of combined PET/CT in recurrent colorectal carcinoma.

When available, PET/CT currently appears the diagnostic tool of choice. In the near future, combined PET/MRI may further enhance the diagnostic algorithm.

9.1 Introduction

Early detection of recurrent colorectal carcinoma has become more important in the past decade, as the treatment options for localized disease have improved significantly. However, aggressive locoregional interventions (e.g. partial liver resections, radiofrequency ablation (RFA) of liver metastases, resections of pulmonary metastases) are considered futile in the presence of metastases elsewhere (1). Therefore, detection of tumor sites throughout the body is needed with high sensitivity and specificity. For further patient management with regard to invasive therapy, accurate information about the local extent of tumor sites is also necessary.

Tumor visualization is traditionally performed using merely anatomical imaging devices such as Computer Tomography (CT), Ultrasound (US), and Magnetic Resonance Imaging (MRI). Functional imaging may be of additional value, especially when anatomical imaging is cumbersome. Visualization of metabolism with ^{18}F-fluor-deoxy-glucose in positron emission tomography (FDG-PET) is a valuable tool for detection of primary and recurrent colorectal cancer (2-4). Tumor sites may be detected throughout the body with high contrast resolution. However, exact localization and demarcation of lesions with PET is hindered by a relatively low spatial resolution, and lack of anatomical reference.

The added value of simultaneous stand-alone FDG-PET and CT has been demonstrated (5). As a next step, the theoretical benefit of joined capabilities of CT (anatomical reference) and FDG-PET (accurate tumor detection) have led to the practice of fusion of images as obtained by PET and by CT. Although promising (6-8), the technique is relatively new and has a limited availability. Furthermore, PET/CT image fusion may suffer from artefacts, and the exact clinical value has not been fully established as yet. Therefore, the role of PET/CT image fusion must be weighed carefully against other, more widely available, modalities.

9.2 Integration of PET and CT

When considering the combination of PET and CT, different methods of fusion are available. The most prevailing approach today is ‘visual fusion’, where two scans are held side-by-side for comparison and correlation. Discrepancies between PET and CT may be cleared with this established technique. When further uncertainties persist, integration of the images can prove to be of additional value. But before attempting to integrate PET and CT images, some specific issues must be considered (9).

Scanning characteristics

Tissues appear differently on PET and on CT images. CT shows anatomy with high spatial resolution, but with low contrast resolution for soft tissues. On the other hand, PET visualizes pathological sites with high contrast resolution, but spatial resolution is limited to the range of 4-7 mm, and surrounding normal anatomical structures are hardly visualized. Due to these characteristics, strong discrepancies may exist between CT and PET images. Benign lesions may appear unequivocal on CT, but may be negative on FDG-PET (e.g. cyst, hemangioma, scar tissue), while intensely FDG-positive lesions may be imperceptible on CT (e.g. local recurrence, liver metastasis). These characteristics complicate visual recognition and correlation.

Positional differences may exist between PET and CT because of repositioning and/or accidental (in)voluntary motion. Organs may be displaced or changed in size (e.g. bowel motion, gastric emptying, bladder filling between PET and CT scanning). Administration of furosemide may contribute to such discrepancies. The main problem is breathing motion. PET is obligatory acquired during free breathing due to the duration of the scanning procedure (20-60 minutes), by default resulting in slightly blurred images in the upper abdomen. For correlation purposes, CT acquisition must be adapted to match these images by scanning during either free breathing or timed unforced expiration (9). Failure to do this correctly will result in serious localization errors, as the diaphragm (including lower lung fields and upper abdominal organs such as the liver) will be relatively displaced.

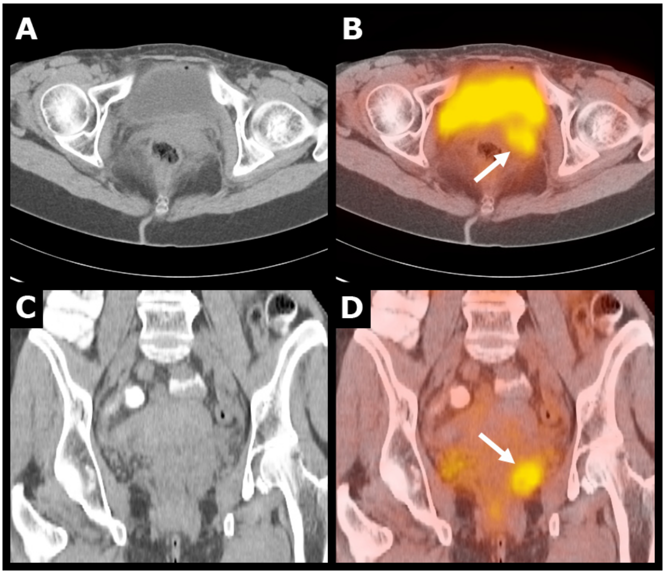


Figure 9.1
Image fusion of contrast-enhanced CT and FDG-PET. Shown are coronal slices of CT (A) and PET (B) through the abdomen. On CT images, the pelvic recurrence is hard to appreciate due to extensive tissue masses. Within these masses, the PET image clearly shows a local recurrence behind the bladder (white arrow). Image fusion with CT provided sufficient anatomical reference to guide a surgical approach.

Software fusion of PET and CT

When separate CT and PET images are available these may be integrated using specialized software. In such procedures certain preconditions need to be met. Identical positioning is a prerequisite. The issue of artefacts due to breathing motions needs to be addressed by breathing instructions. The time gap between scans must be limited, in order to avoid discrepancies due to disease progression (or regression) during the interval. Specific software and operator experience are needed. As a whole, the procedure is lengthy, logistically complex, and it has a serious risk on registration errors. Some authors do report adequate results using software fusion, even in the region of the liver (10), but others strongly dissuade the procedure (11). It must at least be accepted that the bladder and diaphragmatic regions – and probably the whole abdomen – have a limited accuracy in image registration.

Integrated PET/CT scanning

A so-called hybrid scanner consists of in-line placed separate CT and PET scanners, which may acquire both scans consecutively without repositioning the patient. Fusion of images obtained by these two modalities is often referred to as "hardware fusion", although this term ought to be reserved for situations where multiple images are acquired by a single detector system at the same time. "Hardware" PET/CT fusion as currently available reduces (but not fully eliminates) many of the abovementioned positioning problems, but the need for an adequate breathing protocol remains. Other problems such as bladder filling and bowel motion are reduced to

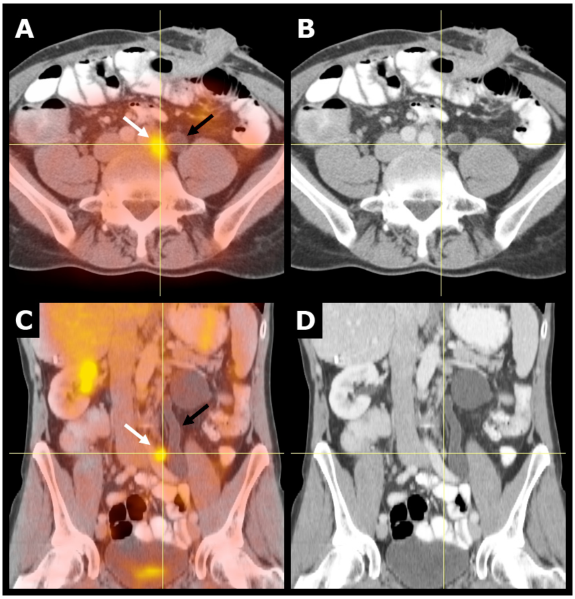


Figure 9.2
Image fusion of contrast-enhanced CT and FDG-PET. The images show transverse slices (A, B) and coronal slices (C, D) through the abdomen of a patient, who previously underwent primary resection of a sigmoid carcinoma. The PET image clearly showed a pathological lesion (white arrow), but the cause remained unclear as no clear lymph node was found, and the dilated ureter suggested another explanation (black arrow). Image fusion with CT could demonstrate correlation with a lymph node that was overlooked before.

acceptable levels. When compensating all abovementioned sources of errors, a fusion error below ~10 mm is generally achievable in the abdomen (12). In specific cases this accuracy may not be reached, for example when a patient is not able to comply with breathing instructions. This source of errors is of high importance when considering the liver, as the result may be misplacement of liver lesions in the lung or vice versa, albeit in a low percentage of scans (13).

When using a hybrid PET/CT scanner, the CT images can be used for attenuation correction of the PET images. Although convenient, as the total scanning time can be reduced by ±35%, any artefact in the CT images may cause secondary artefacts in the PET images. Examples of such artefacts are false-positive hotspots related to attenuating metal such as prosthesis or clips (14), and hotspots related to intravenous/oral contrast (15). Further discrepancies between the PET images and the CT images may result from bowel movements (16), or when the patient accidentally did move between the two scans.

Balancing the benefits

Integration of PET and CT can provide synergistic benefit regardless of the applied technique. Hybrid PET/CT is more expensive than software fusion, but it delivers a fast, logistically easy and more reliable image correlation procedure. A definitive advantage of hybrid PET/CT is that visual fusion and software fusion may be impossible or inadequate when demanded ad-hoc (11). In case of unexpected findings, integrated PET/CT scanning will provide adequate images, while software image fusion is likely to result in suboptimal results.

Interpretation

While fused PET/CT images do appear straightforward, the abovementioned characteristics indicate that the images may not be easy to interpret. The true benefit of integrated PET/CT depends not only on integration of images, but also on the integration of expert opinions. Therefore, it is strongly advised to perform joint reading sessions with the radiologist, nuclear medicine physician and preferably also the referring physician to reach consensus.

9.3 PET/CT in recurrent colorectal carcinoma

In follow-up of colorectal carcinoma, or suspected recurrence (e.g. detectable CEA level, residual or newformed tissues), the clinically relevant questions to be answered include: *Where are potentially malignant tissues localized, is a specific lesion malignant or not, and what is the local extent of a specific lesion?* An important role of imaging is to guide rational use of additional invasive diagnostic procedures (e.g. liver biopsy, colonoscopy, et cetera). A second role is visual demarcation of lesions to guide locoregional therapy. The role of PET/CT in relation to other imaging modalities depends on the indications for the procedure.

Local recurrence

CT is not very accurate for early detection of local recurrence of colorectal carcinoma, due to the distorted local anatomy after operation. Selzner et al. demonstrated a sensitivity of only 53% for CT, and a much better sensitivity for FDG-PET of 93% (7). Such excellent sensitivity in detection of local recurrence also applies in the evaluation after external beam therapy (17). On the other hand, lacking anatomical reference hampers exact localization and evaluation of the extent of local pathology on PET. Since these data are essential when considering therapeutic intervention such as re-excision or irradiation, PET/CT may be of great value. An example of local recurrence detection and localization is provided in figure 9.1. Therefore, for detection and evaluation of local recurrence, it is advised to perform PET/CT when available.

Lymph node metastases

Abdominal lymph node metastases from colorectal carcinoma tend to be small. Many involved lymph nodes are below 1 cm in diameter, explaining the poor sensitivity of CT. Some small metastases can be detected by FDG-PET, be it with a poor sensitivity of 29%, but with a high specificity of 88% (18). Problems arise when a hotspot on PET may correlate with several anatomical structures, activity excreted in the urinary tract, blood vessels, bowel polyps, or physiological bowel uptake. In these cases, PET/CT can adequately identify a hotspot, and settle the diagnosis. Figure 9.2 illustrates PET/CT localization of a pathological lymph node.

Liver metastases

Ruers et al. demonstrated that FDG-PET as a stand-alone modality improves diagnostic work-up in patients with liver metastasis when added to conventional diagnostic imaging. Furthermore, it has an impact on and improves therapeutic management (4). Integrated PET/CT can provide further improved value especially in postoperatively deformed livers with scar tissue and artificial materials, cases where sensitivity and specificity are relatively low for both CT and MRI (19,20). After local ablative therapy, PET may detect recurrence of liver metastasis earlier than CT (3,21), but correlation with CT is needed for more exact localization (8). Conversely, CT may turn false-positive in the rim of the lesions because of hyperperfusion after RFA, while FDG-PET remains reliable (22). MRI using enhancement with manganese containing contrast may further improve detection of liver metastases and provide additional information on the nature of liver lesions (23). Figure 9.3 demonstrates that FDG-PET is not affected by scar tissue and artificial materials. For the detection of liver metastasis after hepatectomy a sensitivity of 100% and specificity of 89% was demonstrated for PET/CT, while the specificity of contrast enhanced CT dropped to 50% for this specific patient category (7). An example of recurrent

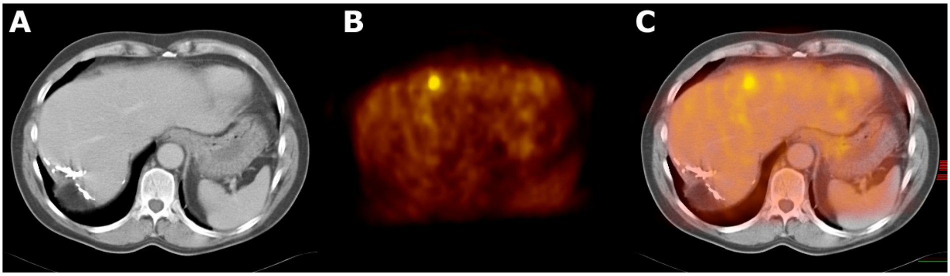


Figure 9.3
Image fusion of contrast-enhanced CT and FDG-PET. The images show transverse slices through the liver, from CT (A), PET (B) and PET/CT (C). PET shows a clear metastasis in the liver that is hardly visible on CT, indicating the high sensitivity of FDG-PET, but also illustrating the need for correlation with anatomical imaging. The image also illustrates that FDG-PET is unaffected by the extensive residual changes and surgical clips in the right liver lobe, after partial liver resection.

metastasis in the liver resection area, not recognized on CT and MRI but detected by FDG-PET and localized by image fusion, is shown in figure 9.4. For evaluation of liver metastases PET/CT appears to be the technique of choice.

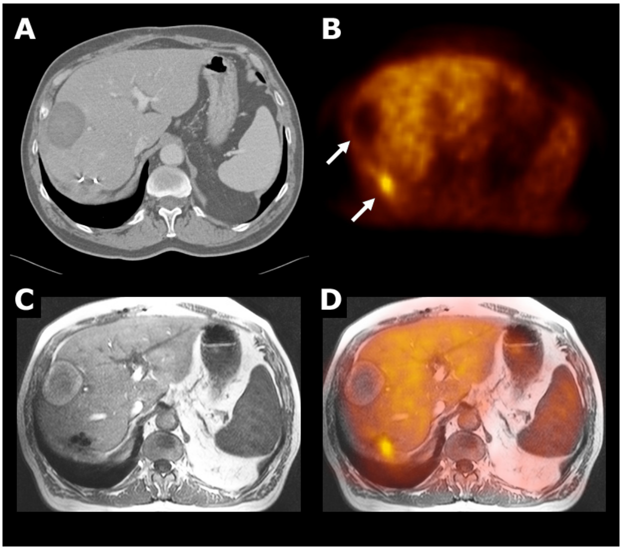
Extrahepatic metastases

The straightforward overview of the whole body in a standard procedure is a major benefit of FDG-PET, thus providing information on extrahepatic metastases. Detection of extrahepatic metastases does have direct impact on patient management. Lai et al. demonstrated that 29% of patients with liver metastases appeared inoperable because FDG-PET detected extrahepatic metastases (2). In recurrent colorectal carcinoma, most extrahepatic distant metastases will be pulmonary metastases. Detection of these metastases is of particular importance as surgical intervention may still be possible, by combining liver surgery with resection of a limited number of pulmonary lesions (24,25). Both CT and FDG-PET have demonstrated high sensitivity for pulmonary lesions, but PET may be particularly helpful in discriminating benign from malignant lesions (26). FDG-PET has also demonstrated added value in detection of other extrahepatic distant metastases such as bone metastases (27). In unexpected extrahepatic lesions detected by PET, exact localization may be very hard without correlative anatomical imaging as provided by PET/CT. This also applies to the detection of unexpected second primaries, which may occur in approximately 1% of cases (28).

Lesion characterization

Regardless of the type of lesion as seen on imaging, differentiation of benign from malignant disease is always a challenge. Both CT and FDG-PET can contribute to the final diagnosis, but

Figure 9.4
Software image fusion of CT, MRI and FDG-PET. The images show transverse slices through the liver of a patient who underwent prior RFA treatment (upper arrow) and resection of liver metastasis of colon carcinoma. Both CT and MRI are difficult to interpret in the region of the surgical clips. The PET image clearly shows a recurrent liver metastasis (lower arrow), which could be localized only after image fusion. This permitted guided locoregional therapy.



combination of both modalities delivers the strongest diagnostic tool (29,30). Given this asset, we consider PET/CT the best option when atypical lesions need to be characterized at the highest possible level of accuracy, especially in cases where a definitive diagnosis through pathology can not be obtained.

9.4 Future developments

The true clinical value of FDG-PET – and the added value of PET/CT scanners – should ideally be clarified by prospective clinical trials. But a true comparison between separately acquired PET and CT images, visual fusion, software fusion, and integrated PET/CT images can hardly be achieved, as this implies the acquisition of multiple scans with a high cumulative radiation burden to the patient. As a result of the rather limited scientific evidence, the current choices for implementation of FDG-PET in diagnostic strategies appear rather random, and large variations exist among institutes. This also applies to the application of hybrid PET/CT scanning for various specific questions. Nevertheless, scientific evidence about the diagnostic values of PET and PET/CT is increasing rapidly, and is eagerly awaited.

New PET tracers

Besides visualization of glucose metabolism with FDG, PET scanning may be applied for in-vivo noninvasive evaluation of other tissue characteristics using tracers other than FDG. For example, DNA synthesis activity may be quantitatively assessed using ^{18}F-fluor-deoxy-thymidine (FLT), reflecting cell proliferation and tumor growth (31). The exact clinical applicability of FLT, as well as of several other tracers currently under investigation, is at present even less clear than the utility of FDG-PET. It is to be expected that many new tracers will accumulate selectively in pathological lesions, and will show poor or no normal tissue activity. These images may therefore be uninterpretable without integration of PET and CT.

Integration of PET and MRI

The combination of PET and CT is not the only possibility, nor is it a perfect solution. On theoretical grounds it is preferable to combine PET with (functional) MRI, for better soft tissue evaluation with a relatively low radiation burden. An excellent example of the application of PET/MRI fusion is accurate delineation of malignant lesions in the liver, to allow optimally guided locoregional therapeutic intervention. The PET/MRI fusion procedure is already possible when using software fusion; an example is shown in figure 9.4. It is expected that integrated PET/MRI scanners will become clinically available in the next five years.

9.5 Conclusions

The combination of PET and CT is currently proving itself as a valuable tool in the diagnostic strategy for detection of recurrent colorectal carcinoma, especially in the field of staging before surgical reinterventions. This has an impact on diagnosis and choice of therapy. In this view, the application of separate PET and CT is not to be considered ‘second class’, when visually correlated adequately. Although unbiased supporting literature is currently limited, hardware integrated PET/CT using a hybrid scanner does seem to be able to improve diagnostic accuracy over correlated stand-alone PET and CT in several specific cases. As software image fusion is prone to error, this technique should be used with caution and should be reserved for specific applications.

The largest benefit from integration of PET and CT images depends on the integration of knowledge. This implies joint consensus reading by a multidisciplinary team. This will be of even greater importance when new PET tracers and new MRI applications enter the clinical field. With the increasing availability of integrated PET/CT scanners, it is to be expected that clinical use and experience will rapidly expand. However, a critical review of indications and added value of these techniques are a prerequisite for rational application and maximum diagnostic yield.

9.6 References

1. Gayowski TJ, Iwatsuki S, Madariaga JR, Selby R, Todo S, Irish W, Starzl TE: Experience in hepatic resection for metastatic colorectal cancer: analysis of clinical and pathologic risk factors. Surgery. 1994;116:703-710
2. Lai DT, Fulham M, Stephen MS, Chu KM, Solomon M, Thompson JF, Sheldon DM, Storey DW: The role of whole-body positron emission tomography with [18F]fluorodeoxyglucose in identifying operable colorectal cancer metastases to the liver. Arch Surg. 1996;131:703-707
3. Langenhoff BS, Oyen WJ, Jager GJ, Strijk SP, Wobbes T, Corstens FH, Ruers TJ: Efficacy of fluorine-18-deoxyglucose positron emission tomography in detecting tumor recurrence after local ablative therapy for liver metastases: a prospective study. J Clin Oncol. 2002;20:4453-4458
4. Ruers TJ, Langenhoff BS, Neeleman N, Jager GJ, Strijk S, Wobbes T, Corstens FH, Oyen WJ: Value of positron emission tomography with [F-18]fluorodeoxy-glucose in patients with colorectal liver metastases: a prospective study. J Clin Oncol. 2002;20:388-395
5. Wiering B, Ruers TJ, Oyen WJ: Role of FDG-PET in the diagnosis and treatment of colorectal liver metastases. Expert Rev Anticancer Ther. 2004;4:607-613
6. Kamel IR, Cohade C, Neyman E, Fishman EK, Wahl RL: Incremental value of CT in PET/CT of patients with colorectal carcinoma. Abdom Imaging. 2004;29:663-668
7. Selzner M, Hany TF, Wildbrett P, McCormack L, Kadry Z, Clavien PA: Does the novel PET/CT imaging modality impact on the treatment of patients with metastatic colorectal cancer of the liver? Ann Surg. 2004;240:1027-1034
8. Veit P, Antoch G, Stergar H, Bockisch A, Forsting M, Kuehl H: Detection of residual tumor after radiofrequency ablation of liver metastasis with dual-modality PET/CT: initial results. Eur Radiol. 2005;
9. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, Corstens FH: PET/CT: panacea, redundancy, or something in between? J Nucl Med. 2004;45 Suppl 1:15S-24S
10. van Dalen JA, Vogel W, Huisman HJ, Oyen WJ, Jager GJ, Karssemeijer N: Accuracy of rigid CT-FDG-PET image registration of the liver. Phys Med Biol. 2004;49:5393-5405
11. Kim JH, Czernin J, len-Auerbach MS, Halpern BS, Fueger BJ, Hecht JR, Ratib O, Phelps ME, Weber WA: Comparison between 18F-FDG PET, in-line PET/CT, and software fusion for restaging of recurrent colorectal cancer. J Nucl Med. 2005;46:587-595
12. Nakamoto Y, Tatsumi M, Cohade C, Osman M, Marshall LT, Wahl RL: Accuracy of image fusion of normal upper abdominal organs visualized with PET/CT. Eur J Nucl Med Mol Imaging. 2003;30:597-602

13. Osman MM, Cohade C, Nakamoto Y, Marshall LT, Leal JP, Wahl RL: Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. J Nucl Med. 2003;44:240-243

14. Goerres GW, Ziegler SI, Burger C, Berthold T, Von Schulthess GK, Buck A: Artifacts at PET and PET/CT caused by metallic hip prosthetic material. Radiology. 2003;226:577-584

15. Dizendorf E, Hany TF, Buck A, Von Schulthess GK, Burger C: Cause and magnitude of the error induced by oral CT contrast agent in CT-based attenuation correction of PET emission studies. J Nucl Med. 2003;44:732-738

16. Nakamoto Y, Chin BB, Cohade C, Osman M, Tatsumi M, Wahl RL: PET/CT: artifacts caused by bowel motion. Nucl Med Commun. 2004;25:221-225

17. Moore HG, Akhurst T, Larson SM, Minsky BD, Mazumdar M, Guillem JG: A case-controlled study of 18-fluorodeoxyglucose positron emission tomography in the detection of pelvic recurrence in previously irradiated rectal cancer patients. J Am Coll Surg. 2003;197:22-28

18. Kantorova I, Lipska L, Belohlavek O, Visokai V, Trubac M, Schneiderova M: Routine (18)F-FDG PET preoperative staging of colorectal cancer: comparison with conventional staging and its impact on treatment decision making. J Nucl Med. 2003;44:1784-1788

19. Dromain C, de BT, Elias D, Kuoch V, Ducreux M, Boige V, Petrow P, Roche A, Sigal R: Hepatic tumors treated with percutaneous radio-frequency ablation: CT and MR imaging follow-up. Radiology. 2002;223:255-262

20. Lim HK, Choi D, Lee WJ, Kim SH, Lee SJ, Jang HJ, Lee JH, Lim JH, Choo IW: Hepatocellular carcinoma treated with percutaneous radio-frequency ablation: evaluation with follow-up multiphase helical CT. Radiology. 2001;221:447-454

21. Blokhuis TJ, van der Schaaf MC, van den Tol MP, Comans EF, Manoliu RA, van dS, Jr.: Results of radio frequency ablation of primary and secondary liver tumors: long-term follow-up with computed tomography and positron emission tomography-18F-deoxyfluoroglucose scanning. Scand J Gastroenterol Suppl. 2004;93-97

22. Antoch G, Vogt FM, Veit P, Freudenberg LS, Blechschmid N, Dirsch O, Bockisch A, Forsting M, Debatin JF, Kuehl H: Assessment of liver tissue after radiofrequency ablation: findings with different imaging procedures. J Nucl Med. 2005;46:520-525

23. Reimer P, Schneider G, Schima W: Hepatobiliary contrast agents for contrast-enhanced MRI of the liver: properties, clinical development and applications. Eur Radiol. 2004;14:559-578

24. King J, Glenn D, Clark W, Zhao J, Steinke K, Clingan P, Morris DL: Percutaneous radiofrequency ablation of pulmonary metastases in patients with colorectal cancer. Br J Surg. 2004;91:217-223

25. Steinke K, Glenn D, King J, Clark W, Zhao J, Clingan P, Morris DL: Percutaneous imaging-guided radiofrequency ablation in patients with colorectal pulmonary metastases: 1-year follow-up. Ann Surg Oncol. 2004;11:207-212

26. Lowe VJ, Fletcher JW, Gobar L, Lawson M, Kirchner P, Valk P, Karis J, Hubner K, Delbeke D, Heiberg EV, Patz EF, Coleman RE: Prospective investigation of positron emission tomography in lung nodules. J Clin Oncol. 1998;16:1075-1084

27. Bohdiewicz PJ, Wong CY, Kondas D, Gaskill M, Dworkin HJ: High predictive value of F-18 FDG PET patterns of the spine for metastases or benign lesions with good agreement between readers. Clin Nucl Med. 2003;28:966-970

28. Ishimori T, Patel PV, Wahl RL: Detection of unexpected additional primary malignancies with PET/CT. J Nucl Med. 2005;46:752-757

29. Antoch G, Freudenberg LS, Beyer T, Bockisch A, Debatin JF: To enhance or not to enhance? 18F-FDG and CT contrast agents in dual-modality 18F-FDG PET/CT. J Nucl Med. 2004;45 Suppl 1:56S-65S

30. Cohade C, Osman M, Leal J, Wahl RL: Direct comparison of (18)F-FDG PET and PET/CT in patients with colorectal carcinoma. J Nucl Med. 2003;44:1797-1803

31. Francis DL, Freeman A, Visvikis D, Costa DC, Luthra SK, Novelli M, Taylor I, Ell PJ: In vivo imaging of cellular proliferation in colorectal cancer using positron emission tomography. Gut. 2003;52:1602-1606

10

Accuracy of rigid CT – FDG-PET image registration of the liver

Jorn A. van Dalen [1,2]
Wouter V. Vogel [1]
HenkJan Huisman [2]
Wim J. G. Oyen [1]
Gerrit J. Jager [2]
Nico Karssemeijer [2]

Departments of Nuclear medicine [1] and Radiology [2],
Radboud University Nijmegen Medical Centre, the Netherlands

Physics in Medicine and Biology, 2004
Volume 49, pages 5393–5405.

Abstract

Diagnostic and surgical strategies could benefit from accurate localization of liver malignancies via CT–FDG-PET image registration. However, registration uncertainty occurs due to protocol differences in data-acquisition, the limited spatial resolution of positron emission tomography (PET) and the low uptake of ^{18}F-fluor-deoxy-glucose (FDG) in normal liver tissue. To assess this uncertainty, methods were presented to estimate registration precision and systematic bias.

Methods: A semi-automatic, organ-focused method was investigated to minimize the uncertainty well beyond the typical uncertainty of 5–10 mm obtained by commonly available methods. By restricting registration to the liver region and by isolating the liver on computed tomography (CT) from surrounding structures using a thresholding technique, registration was achieved using the mutual information-based method as implemented in insight toolkit (ITK). CT and FDG-PET images of 10 patients with liver metastases were registered rigidly a number of times. Results of the organ-focused method were compared to results of three commonly available methods (a manual, a landmark-based and a 'standard' mutual information-based method) where no dedicated image processing was performed.

Results: The proposed method outperformed the other methods with a precision (mean ± S.D.) of 2.5±1.3 mm and a bias of 1.9 mm with a 95% CI of [1.0, 2.8] mm.

Conclusion: Unlike the commonly available methods, our approach allows for robust CT–FDG-PET registration of the liver, with an accuracy better than the spatial resolution of the PET scanner that was used.

10.1 Introduction

Detection of malignancies in the liver using computed tomography (CT) is difficult when tumor tissue has an electron density similar to normal liver tissue. Even when contrast-enhancing techniques are used, the detection of liver malignancies can be difficult. Particularly in a deformed liver after therapeutic interventions, a CT image can be uninterpretable due to the presence of scar tissue, necrotic remnants and artificial materials as staples. Positron emission tomography (PET) with ^{18}F-fluor-deoxy-glucose (FDG) visualizes glucose metabolism, rather than anatomy, making it suitable to detect tumor tissue. Often, localizing FDG-PET activity can be a problem and therefore integrating FDG-PET and CT data allows correlation of pathological FDG activity to CT-detected anatomical structures. This may improve the diagnosis, enabling better therapeutic management by clinicians (1,2).

Integration of FDG-PET and CT images consists of two steps. The first step is the registration, i.e., the determination of the geometrical transformation of one of the acquired images needed to fit the other image. The second step of the integration is the fusion, required for the integrated display of the data. This is mainly important for adequate visualization.

Registration problems occur due to the limited spatial resolution of PET and the moderate uptake of FDG in normal liver tissue. Furthermore, due to the time difference and differences in the protocols in the data-acquisition of the PET and CT data sets, differences occur in the position, orientation and shape of the liver, which also make the registration difficult. The use of a combined PET-CT scanner does not necessarily provide a satisfactory solution since respiration and patient motion induce artefacts on CT (3-7) and thus on the fused PET-CT images.

Essential for using image integration in the clinic is knowing the registration accuracy. Accuracy measures how close to the ‘true’ value a measurement lies. It includes both precision, which measures how closely results can be duplicated, and systematic bias, or, i.e., the combination of random and systematic uncertainties. Quantification of the registration accuracy is non-trivial for the reason that the ‘ground truth’ is lacking (8). However, transferring uncertainty measures obtained by reference, statements can be made related to the likelihood for the uncertainty to exceed a certain bound. In literature, the amount of registration-validation studies is small compared to studies involving image registration. Fitzpatrick et al. have provided a critical and direct measure of the registration accuracy for rigid-body, point-based registration (9). It is the only measure that expresses the displacement between any two corresponding points in the images. Unfortunately, lack of (anatomical) landmarks on both FDG-PET and CT images prevents us to use this validation measure.

In this paper, methods are presented to estimate the precision and systematic bias of CT–FDG-PET liver registration. Precision and bias are then compared for four methods. One method focuses on the liver by including image processing before performing the registration. The other three methods are conventionally available techniques where no such pre-processing is performed. The organ-specific approach is hypothesized to give more accurate registration results.

10.2 Materials

Images of 10 consecutive patients with histologically confirmed colorectal carcinoma with liver metastases were used. All patients have a prior history of resection of the primary tumor and at least once underwent surgical treatment for liver metastases. Each patient received a CT and FDG-PET scan within a week.

CT was performed using a multi-slice spiral CT scanner (Siemens Somatom Volume Zoom, Erlangen, Germany) with specifications as given in table 10.1. After contrast injection (90 ml Omnipaque-350, or 100 ml Xenetix-300), a scan of the whole liver was taken during unforced expiration breath-hold. This protocol ensures optimal positioning of the diaphragm and liver for image registration with PET (5). Therefore, using this protocol, differences in the shape of the liver between CT and PET are expected to be minimal, allowing rigid registration. Furthermore, the arms of the patient were above the head, for better image quality.

PET scans were acquired using a full-ring dedicated PET scanner (ECAT Exact 47, Siemens/CTI, Knoxville, Tennessee, USA). Patients fasted for a minimum of 6 h before intravenous injection of the radiopharmaceutical ^{18}F-fluor-deoxy-glucose (FDG). The PET protocol consisted of an emission scan from the hips to the base of the skull, with ^{68}Ge-based attenuation correction. The arms of the patient were next to the body. Due to the long PET scanning time of typically 1 h it would be inconvenient for the patient to put the arms above the head. Specifications of the FDG-PET scan are given in table 10.1.

Modality	**Specifications**				
	PS (mm^2)	ST (mm)	FOV (mm)	V (kVp)	R (mAs)
CT	0.74 – 0.78	6 – 8	380 – 400	120	150
	PS (mm^2)	ST (mm)	FOV (mm)	Resolution (mm)	A (MBq)
FDG-PET	5.15	5.15	562	7	250

Table 10.1
Specifications of the CT and FDG-PET scans. For both types the pixel size (PS) in the axial view, slice thickness (ST) in the axial direction and field of view (FOV) are given. For the CT scan, the x-ray high voltage peak (V) and the dose rate (R) are given. For the FDG-PET scan the spatial resolution (full width at half maximum) and the initial radioactivity (A) are given.

10.3 Methods

A registration method was investigated to minimize uncertainties due to position and orientation differences of the liver between the FDG-PET and CT images. To achieve minimum uncertainties, a mutual information-based registration method was applied onto image volumes of interest focused on the liver. Within these volumes, image processing prior to the registration was performed to further isolate the liver on the CT from surrounding structures. This organ-focused mutual information-based method (OFMI) was compared to three commonly available methods: a manual method (MAN), a landmark-based method (LM) and a 'standard' mutual information-based method (SMI) as described by Wells et al (10). Using comparable breathing protocols for CT and PET, differences in the shape of the liver in both image modalities were expected to be minimal. Hence, the registration transformation T was taken rigid, i.e., consisting of three rotation and three translation parameters.

Registration

The mutual information (MI) similarity measure was introduced for matching of medical images in 1995 (11,12). Being a voxel-based registration method, the mutual information-based registration method optimizes a functional measuring the similarity of all geometrically corresponding voxel pairs for some feature. Consequently, the achievable accuracy of this registration method is not limited by the voxel size of the images (13), or rather (in this study) by the spatial resolution of the PET scanner.

The MI metric implementation follows the method as specified by Viola and Wells (12,14). In this implementation, probability densities are estimated from the image data using the Parzen-Window scheme (15). In particular, N_S and N_R elements of two samples, $S_i \in S$ and $R_i \in R$, are drawn from the image: the first to compute the density, the second to estimate the entropy. The kernel function was a Gaussian with a width ψ. Having estimated entropy, an approximation can be found, $\widehat{\frac{\mathrm{d}MI}{\mathrm{d}T}}$, for the derivative of the MI similarity measure, with respect to T (12,14). The stochastic-gradient descent scheme was used to find optimal transformation parameters:

$$T_{j+1} = T_j + \lambda \frac{\mathrm{d}\widehat{MI}}{\mathrm{d}T}\bigg|_{T_j} \qquad \text{Equation (10.1)}$$

where T_j indicates the values of the transformation parameters at iteration j and λ is the learning rate. Different learning rates were used for translations and rotations. The mutual information-based method described above is available in ITK (16).

Before the method was applied, the user first aligned the FDG-PET and CT images manually. The manual registration result was used as the initial input in the MI optimalization process, i.e., T_0 in equation 10.1.

By restricting the registration to the region of the liver, surrounding structures, such as the heart, cannot disturb the registration. Therefore, user-identified boundaries of the liver as seen on the CT image were used for creating a volume of interest (VOI) that contains the liver. As indicated in figure 10.1, the boundaries were constructed based on four selected points on the liver: the point with the highest axial (z) coordinate (top point), with the lowest axial coordinate (bottom point), with the lowest coronal (y) coordinate (front point) and with the lowest sagittal (x) coordinate (right point). These four points define four planes (orthogonal to either the x, y or z direction) of a box. The remaining two planes were constructed using the right point translated over a distance ΔX in the x direction and the front point translated over a distance ΔY in the y direction. The values for ΔX and ΔY were chosen such that surrounding structures of the liver (e.g., the heart) hardly take part in the registration. This way, it can happen that the box does not contain the liver completely. This is not alarming, since the potentially missing parts of the liver can hardly be isolated from surrounding structures on the FDG-PET image, so that it cannot give a significant contribution to the registration performance. However, it is most important that the sharply defined edges in the right liver lobe (around the top, bottom, front and right points) are within the VOI that is used for the registration. Therefore, the VOI was determined by increasing the constructed box in all directions by 5 mm. Figure 10.1 illustrates the determination of the VOI. Once a VOI was created on the CT image, automatically an overlapping VOI on the FDG-PET image was created.

It was found that the MI similarity can be increased by first thresholding the CT image. The typical CT Hounsfield number for the liver is 1050 HU (17). Setting all grey values outside the region [1000, 1200] HU to 1000, the liver was well isolated from surrounding structures. As an illustration, figure 10.2.A shows an example of the image pre-processing to arrive at the VOIs that are used for mutual information-based registration. Figures 10.2.B and 10.2.C show examples of fusion results.

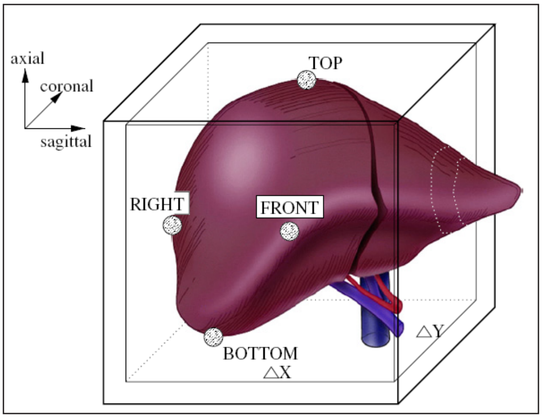


Figure 10.1
Determination of the VOI, based on a box that is constructed using four liver points identified by the user (see the text).

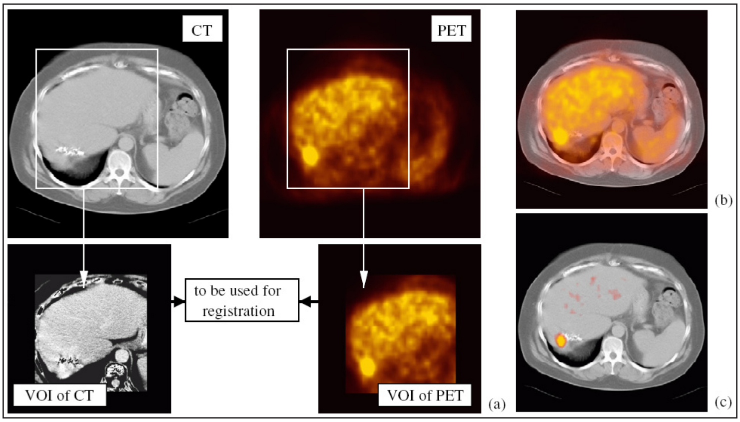


Figure 10.2
(A) Example of the image pre-processing to arrive at the VOIs that are used for mutual information-based registration. An axial view is shown of the original CT and FDG-PET images. From these images VOIs are determined. In addition, for the CT image a thresholding is performed; (B) a fusion example after registration; (C) a fusion result using different window and level settings for FDG-PET, showing clearly the location of a lesion, which was hardly visible on CT alone due to scar tissue and staples.

Prior to the mutual information-based registration of the FDG-PET and the thresholded CT images of the liver, the statistical distributions of the images' grey levels were normalized such that the grey values have zero mean and unit variance. Moreover, the CT image was resampled so that its original voxel size was reduced by a factor 6 in both the x and y directions. This way, the voxel dimensions were more isotropic and comparable to the FDG-PET voxel dimensions, increasing the robustness of the mutual information-based method (10). A tri-linear interpolation scheme was applied onto the images to estimate the image intensity at any possible position in the images.

In this paper, a comparison is made with the results of three commonly available registration methods. These three methods are the following:

- *The manual (MAN) method, i.e., manually aligning the images in 3D*. The operator used an interface where three variable sections of orthogonal planes (x, y and z) through the two images were displayed simultaneously. Hence, by performing registration in only one section at the time, the 3D registration problem was reduced to a series of 2D problems.

- *The landmark-based (LM) method.* Firstly, the user identified on both the CT and FDG-PET images at least four points on the liver: the top, bottom, front and right points. Although difficult, these points could be identified on both images. If other anatomical points on the liver could be identified the user was allowed to do this. Next, registration was performed by matching the identified points in a least square sense. In general, the use of anatomical or artificial orientation points in both images can be of benefit for the registration (8).

- *The 'standard' mutual information-based (SMI) method.* Registration was performed by applying the mutual information-based method (also using the Parzen-Window scheme) to the whole CT image and the PET transmission image that is used for attenuation correction. No image pre-processing (e.g., thresholding on CT) was performed on the images. Further, the registration transformation derived from the PET transmission and CT images was used for fusing the emission FDG-PET and CT images. In some cases the use of the PET transmission image, instead of the emission image, is of benefit for the registration: in general, emission scans contain too little anatomical information to apply standard image registration methods (18-20). Apart from a few organs such as the liver (but also, e.g., the brain, heart and kidneys), the emission scan hardly visualises other anatomy. Sometimes also a combination of emission and transmission PET volumes is used for registration purposes (21,22). We also studied registration of the CT with a constructed PET image consisting of the sum of the emission image and 25% (22) of the transmission image. The manual registration result was used to initialize the MI optimalization process.

Validation

The validation includes precision measurements, measurements of a possible systematic bias of a method, and a visual judgement. The combination of precision and bias of a method reveals the accuracy of a registration. By studying results for different patients, performed by different operators, insight is provided in the accuracy of a registration for a particular patient, performed by a particular operator.

Precision

For each of the four methods (OFMI, MAN, LM and SMI), each CT–FDG-PET pair was registered by four operators: one nuclear medicine physician and three physicists. To obtain the intra-operator variability, the individual operator repeated the registration, for each patient and each method, up to four times. Variability in the registration between different operators was used to determine the inter-operator variability.

To compare the registration results, each computed registration transformation, M, was applied onto four points, $\vec{t}\,'$, $\vec{b}\,'$, $\vec{f}\,'$ and $\vec{r}\,'$, in the FDG-PET image. The points, which were determined once for each patient by a nuclear medicine physician, are the top, bottom, front and right points of the liver, respectively. Next, inter- and intra-operator variabilities of the resulting points, $\vec{t}\,' = M\vec{t}\,'$, $\vec{b}\,' = M\vec{b}\,'$, $\vec{f}\,' = M\vec{f}\,'$ and $\vec{r}\,' = M\vec{r}\,'$ were determined. To assess the registration precision, the points $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$ cannot directly be compared between patients.

Therefore, the average location of each point ($\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$) was computed for each patient and each method. For a given point, the average location was subtracted for each patient and each method. As the number of registrations per patient per method was rather low (in the order of 7), a trimmed (25%) mean method was used to determine a least-biased average. From the resulting vector differences, the lengths were computed, representing the absolute distances from the average points of $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$, denoted by D_t, D_b, D_f, and D_r, respectively. The squared length of these distances follow a χ^2 distribution (23), which can be characterized by a mean and a standard deviation (S.D.). The S.D. indicates variations due to different patients and operators. Again robust estimates were used: the mean was estimated using a trimmed (25%) mean method. The S.D. was estimated using the median of the absolute deviation (MAD) from the median. For all distance distributions, the mean values resulting from the four methods were compared using the Wilcoxon rank sum test (or, the Mann–Whitney test). Furthermore, the analysis of variance (ANOVA) test was used to analyse the influence of different components (intra- and inter-operator components and possible other components).

Systematic bias

To test for systematic differences between the output of the different registration methods, the non-parametric Wilcoxon signed rank test was applied onto paired data points of $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$, resulting from any two registration methods. Here, average locations of $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$ were not subtracted from the individual point locations. In particular, using the approximation that the MAN results are symmetric around the 'ground truth', we estimated systematic uncertainties for the OFMI, LM and SMI methods.

Since visual assessment is likely to provide the test quality assurance of multi-modality registration accuracy (24), the MAN results are assumed to provide the 'ground truth' registration. Also, values of rotation parameters in the registration transformation were analysed to search for biases. If a particular method results in a systematic rotation, this indicates a bias. For each method, rotation distributions were determined for all registrations that were performed and the mean and S.D. were determined. Necessary to support the assertion that the MAN results are symmetric around the 'ground truth' is to test the symmetry properties of the MAN results. Therefore, differences in the MAN distributions from the average for each patient, each point and each axis were computed. Single histograms were determined for each of x, y and z over all points, patients and operators. From these distributions the skewness was computed indicating if the histograms are symmetric about the average value.

Visual judgement

Validation was further performed by visual judgement of the fusion results. For given CT–FDG-PET pairs, operators had to rank the fusion results, corresponding to four different anonymized methods. The sign-test was used to test the hypothesis that one method is not preferred to another method.

10.4 Results

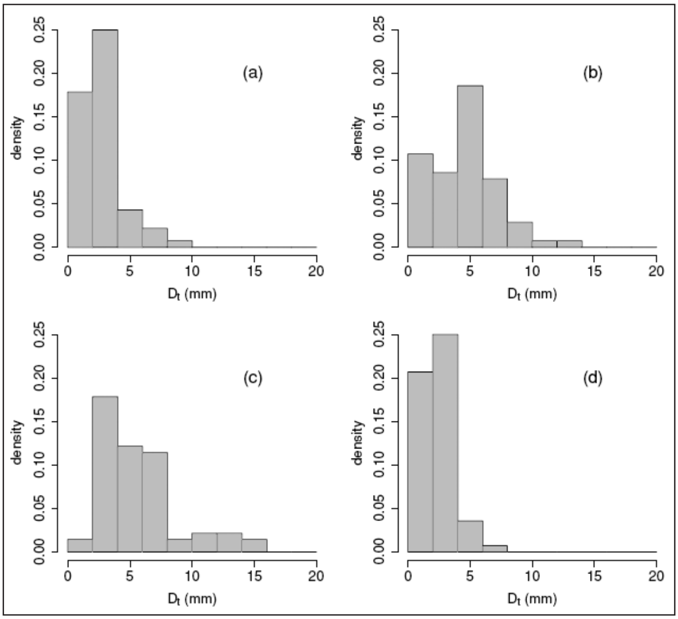
The operators performed 280 (70 for each of the four methods) registrations. Furthermore, experience of Viola and Wells (12,14) and Wells et al (10) and fine tuning resulted in the following parameter settings for the mutual information-based method: ΔX = 160 mm, ΔY = 180 mm, $N_S = N_R$ = 50, for rotations $\lambda = 0.0005$, for translations $\lambda = 0.05$, number of iterations = 20,000 (enough to reach convergence of the transformation parameters) and $\psi = 2$ (width of Parzen Window). Experimentally, the values for ΔX and ΔY were found to be good values to minimize the presence of surrounding structures of the liver, without removing sharply defined edges of the right liver lobe. The total run time for the mutual information-based method, on a PIII/933MHz/128MB, is about 3 min, comparable with the times given in Wells et al. (10). An experienced operator was able to perform the pre-processing within 2 min, so that the total registration time for the OFMI or SMI method was performed within 5 min. The registration time for the MAN or LM method was typically 2 min.

Table 10.2 summarizes the D_t, D_b, D_f, and D_r distributions for each of the four methods. Since the results are highly correlated for a given method, also the average of the means and S.D.s of these points are given. The results of the SMI method were found to be similar to those of the variant of the SMI method where a linear combination of the emission and transmission PET scan was used in the registration. Further, for none of the distance distributions a significant difference in mean distance between the OFMI and SMI results was observed. However, the MAN and LM methods gave a significantly (p < 0.001 for each point) increased mean distance compared to the OFMI and SMI methods.

Method	D_t (mm)		D_b (mm)		D_f (mm)		D_r (mm)		Average	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
OFMI	2.4	1.1	2.4	1.4	2.6	1.5	2.4	1.2	2.5	1.3
MAN	4.4	2.9	4.7	3.0	4.7	2.6	4.7	2.9	4.6	2.9
LM	5.2	2.8	5.0	2.6	6.1	3.2	5.2	3.1	5.4	2.9
SMI	2.2	1.1	2.7	1.5	2.5	1.3	2.7	1.5	2.5	1.3

Table 10.2
Mean and s.d. of the D_t, D_b, D_f, and D_r distributions, for the OFMI, MAN, LM and SMI methods. Also the average over the four distributions is given.

Figure 10.3
Density distributions of the distance D_t, for the methods OFMI (A), MAN (B), LM (C), and SMI (D).



As an illustration, figure 10.3 shows the distance distributions for the top point, for each method. For the other points, similar distributions were obtained. It is clear that the best precision measurements were obtained with the OFMI and SMI methods, without giving big outliers, which makes these methods robust.

For the MAN method none of the operators made use of all six transformation parameters. Only the three translation parameters were used and sometimes one rotation parameter. On average the number of degrees of freedom for the MAN method was found to be 3.9. For the other methods all six parameters were used. In general, the (trimmed) mean of the distance distribution will be smaller when less transformation parameters (less degrees of freedom) are used. Therefore, the means of the MAN method in table 10.2 may be underestimated in comparison to the results of the other methods.

The ANOVA test suggests that, for each method and each point (top, bottom, front, right) there is no statistical significant difference in mean registration results obtained by different operators, nor in mean results when registrations were repeated by the operators. Moreover, the contribution to the total variability originating from either the intra- or interoperator components or possible other components was statistically not significant (for each component $p > 0.05$).

Coordinate	OFMI-MAN (mm)		OFMI-LM (mm)		OFMI-SMI (mm)	
	Difference	95% CI	Difference	95% CI	Difference	95% CI
t_x	-2.5 *	[-3.6, -1.6]	-4.5 *	[-5.6, -3.2]	-7.0 *	[-8.2, -5.7]
t_y	0.4	[-1.1, 1.9]	-4.4 *	[-6.1, -2.8]	-0.8	[-2.6, 0.6]
t_z	0.6	[-0.2, 1.5]	0.9	[-0.1,1.8]	7.0 *	[5.7, 8.3]
b_x	-1.0	[-2.1, 0.1]	-0.9	[-2.0, 0.2]	-3.6 *	[-5.0, -2.1]
b_y	0.1	[-1.0, 1.2]	-3.0 *	[-4.7, -1.4]	-0.7	[-1.7, 0.2]
b_z	0.3	[-0.5, 1.0]	0.1	[-1.2, 0.9]	6.3 *	[5.0, 7.4]
f_x	-2.4 *	[-3.2, -1.6]	-0.8	[-1.9, 0.3]	-6.1 *	[-7.1, -5.0]
f_y	0.4	[-0.9, 1.5]	-2.0 *	[-3.6, -0.7]	-2.3 *	[-3.8, -0.7]
f_z	1.6	[0.0, 3.0]	1.3	[-0.4, 2.7]	9.1 *	[7.0, 11.0]
r_x	-1.7 *	[-2.4, -1.0]	-2.6 *	[-3.4, -1.7]	-5.6 *	[-6.6, -4.4]
r_y	0.3	[-1.0, 1.7]	-3.9 *	[-5.5, -2.6]	-0.3	[-2.0, 1.1]
r_z	0.2	[-0.6, 0.9]	-0.4	[-1.5, 0.7]	6.1 *	[4.9, 7.1]

Table 10.3

Results of the Wilcoxon signed rank test applied onto paired data points of $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$, *for different combinations of two methods. Mean differences of paired data between different methods and a confidence interval (CI) of 95% are given for each coordinate. A '*' sign indicates that the difference is significantly ($p < 0.05$) different from zero.*

Applying the Wilcoxon signed rank test onto paired data points of $\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$, resulting from any possible combination of two methods, we observe significant deviations from zero for mean differences of paired data for x, y and z components. Table 10.3 summarizes the mean differences with a confidence interval of 95%. For each point, we observe that the difference of paired data between the OFMI and MAN results has a mean value consistent with zero for the y and z components, but a significant deviation from zero, up to 2.5 mm, for the x

Coordinate	Systematic uncertainty including 95% CI (mm)					
	OFMI		LM		SMI	
x	1.9 *	[1.0, 2.8]	-0.3	[-1.3, 0.8]	-3.7 *	[-4.8, -2.4]
y	-0.3	[-1.6, 1.0]	-3.6 *	[-5.3, -2.2]	-1.3	[-2.8, 0.0]
z	-0.7	[-1.6, 0.3]	-0.2	[-1.5, 0.9]	6.5 *	[5.0, 7.8]

Table 10.4

Systematic registration uncertainty for the OFMI, LM and SMI methods, including the 95% CI. A '' sign indicates that the uncertainty is significantly ($p < 0.05$) different from zero.*

component. Comparison between results of the OFMI and the LM or SMI method shows much bigger deviations, in all components: up to 4.5 mm when comparing with the LM results and up to 9.1 mm when comparing with the SMI results. Further, we observe that the results for the different points ($\vec{t}$, $\vec{b}$, $\vec{f}$ and $\vec{r}$) are highly correlated.

Assuming the MAN results are symmetric around the 'ground truth', we estimated the systematic uncertainty for the OFMI, LM and SMI methods using the results given in table 10.3. Firstly, mean differences (and CIs) in paired data of the MAN results and any of the OFMI, LM and SMI results were determined, for all four points. Next, the average over the four points of these mean differences and CIs were computed. Since no significant dependence on the position on the liver was observed (see also table 10.3), these average values, summarized in table 10.4, serve as an estimation of the systematic registration uncertainty for the OFMI, LM and SMI methods on any point on the liver. It is clear that the OFMI method yields the smallest bias.

Table 10.5 shows the mean and S.D. of the rotation distributions (around the three axes) of all registrations that were performed. In general, we observe that the mean results are consistent with zero, for all methods. The spread in rotations is typically 1–2°, except for the manual method where it is smaller. This makes sense when realizing that for the manual method, rotations were significantly less applied than for the other methods. The skewness for the single MAN histograms of each of x, y and z over all points, patients and operators was found to be -0.25, -0.06 and 0.25, respectively. Hence, these histograms (each having 70 registrations × 4 points = 280 entries) can be considered symmetrical about their average values, with $p>0.05$. Visual judgement showed that in most cases (60%) the OFMI method was found to be the best. It was found to be significantly ($p<0.05$) better than one of the other methods. Furthermore, the tendency is that the MAN method is preferred to the LM method and the LM method is preferred to the SMI method.

Method	**Rotation x-axis (°)**		**Rotation y-axis (°)**		**Rotation z-axis (°)**	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
OFMI	0.2	1.0	-0.1	1.5	-0.5	1.8
MAN	0.5	0.5	-0.4	0.7	0.4	1.0
LM	-0.6	2.0	0.4	1.2	-1.0 *	1.7
SMI	-0.2	1.3	0.4	1.3	0.2	1.8

Table 10.5
Mean and s.d. of the rotations, for each method, corresponding to all registrations that are performed. A '' sign indicates that the mean is significantly ($p < 0.05$) different from zero.*

10.5 Discussion

In this paper, methods were presented to estimate precision and systematic bias of registrations. Further, we showed that by focusing on the organ of interest (OFMI method), better registration results were obtained than for other conventional methods.

In the validation process we assumed the MAN results to be symmetric around the 'ground truth'. One may have doubts about the negligibility of a systematic bias of the MAN method. In particular, doubts can arise after the observation that the operators did not use all degrees of freedom when performing manual registrations: rotation parameters were hardly used. However, the average performance of the MAN method is based on averaging over both the patients *and* the operators. We expect that most manually induced biases that might be present when looking at a single patient or a single operator cancel each other when averaging over a set of patients *and* various operators. At least, biases due to rotation effects can be neglected. The mean rotation (around the three axes) is found to be consistent with zero, not only for the MAN results, but also for the results of the other methods. Further, the single MAN histograms of each of x, y and z over all points, patients and operators were found to be symmetrical about their average values. This is a necessary observation for the assertion that the MAN results are symmetric around the 'gold standard'. Moreover, we visually observed the tendency that the OFMI method is the best, followed by the MAN method, the LM method and finally the SMI method. Hence, a possible bias of the OFMI method is more likely to be related to differences with the results of the MAN method than to differences with the results of the LM or SMI method. Based on the arguments above, we do not expect that the MAN method is strongly biased. This supports that the estimated systematic uncertainties of the OFMI method, that go up to 1.9 mm (table 10.4), are reasonable and not an underestimation.

Questions may arise as to why the OFMI method outperforms the other methods. Firstly, since the MAN and LM results greatly depend on the operator, the precision of these methods is worse than that of the OFMI and SMI methods (table 10.2). Moreover, the relatively high mean values for the LM method given in table 10.2 are also due to the absence of clear anatomical landmarks on the liver. Next, the SMI method resulted in the biggest systematic uncertainties. There are two reasons why the SMI method fails. Firstly, the SMI method makes use of the whole CT of the abdomen. Hence, registration uncertainties can occur due to position and orientation differences of the liver between the FDG-PET and CT images. Secondly, lack of contrast in transmission PET images prohibits its use for distinguishing between soft tissue. Therefore, by using the transmission PET image, the registration is effectively based on the contours of the body. Due to protocol differences in PET and CT data-acquisition, e.g., differences in the position of the arms, differences occur in the contours of the body, affecting the registration performance in a negative way. As was explicitly checked, failure of the SMI method is not a consequence of not thresholding the CT.

A shortcoming of the OFMI method is that it does not account for deformation differences of the liver between the FDG-PET and CT images. The unforced expiration breath-hold during the CT scan cannot be quantified in an absolute manner. Hence, variations may occur around the optimal unforced expiration breath-hold for registration. As previously observed (5), the unforced expiration breath-hold protocol results in differences in location (0.4 mm (mean), 11.7 mm (1 S.D.)) of the diaphragmatic dome on CT as compared with the diaphragmatic dome on the PET scan. Hence, differences in movement can also be expected when observing the liver. Part of this liver movement will be rigid and part of it will be non-rigid. The non-rigid part is not taken into account in the OFMI method. Using respiratory gating provides insight in the non-rigid liver deformation and hence provides a technique to deal with the non-rigidities (25). Other potential sources of residual deformations can be attributed to the time difference between the data-acquisition of the two images and to other differences between the CT and PET acquisition protocols. Use of non-rigid registrations possibly can improve the registration results. However, a problem in non-rigid registrations is the introduction of new uncertainties due to the necessary segmentation and modelling (8).

The introduction of combined PET–CT systems solves a number of registration problems (26), but it does not solve all. Apart from respiratory-induced artefacts leading to registration uncertainties (3-7), information from other modalities (such as magnetic resonance imaging) cannot be included without an image-integration software application. Therefore, the necessity to develop methods for integrating CT images (or images from another modality) with images of a stand-alone PET will remain.

10.6 Conclusion

We have presented methods to quantitatively compare the registration precision and systematic bias of different registration methods. Measuring both precision and possible bias of a method provides an estimate of the accuracy of a registration. An organ-focused mutual information based method was demonstrated to outperform three other conventional techniques. In particular, the precision (mean ± S.D.) for this method was found to be 2.5 ± 1.3 mm. The systematic uncertainty was found to be 1.9 mm with a 95% CI of [1.0, 2.8] mm. By analysing precision and bias on several points on the liver and observing high correlation between the results, the above-mentioned results are approximately valid to any point on the liver.

Hence, unlike commonly available registration methods, the organ-focused method results in registrations with an accuracy better than the spatial resolution of the PET scanner that was used. Further, the strength of the organ-focused method is that it potentially can be applied on demand on any area it is needed for. In general, it is not clinically interesting to fuse the whole images: it is the PET-positives that cannot visually be located on CT that are interesting. This is only in part of the patient's population and only in part of the lesion case. Hence, we think this method is acceptable for use in the clinic.

10.7 References

1. Cohade C and Wahl RL. Applications of positron emission tomography/computed tomography image fusion in clinical positron emission tomography-clinical use, interpretation methods, diagnostic improvements. Semin. Nucl. Med. 2003:33;228–237
2. Hosten N, Kreissig R, Puls R, Amthauer H, Beier J, Rohlfing T, Stroszczynski C, Herbel A, Lemke AJ and Felix R. Fusion of CT and PET data: methods and clinical relevance for planning laser-induced thermotherapy of liver metastases. Rofo. Fortschr. Geb. Rontgenstr. Neuen Bildgeb. Verfahr. 2002;172:630–635
3. Beyer T, Antoch G, Blodgett T, Freudenberg LF, Akhurst T and Mueller S. Dual-modality PET/CT imaging: the effect of respiratory motion on combined image quality in clinical oncology. Eur. J. Nucl. Med. Mol. Imaging. 2003;30:588–596
4. Chin BB, Nakamoto Y, Kraitchman DL, Marshall L and Wahl R. PET–CT evaluation of 2-deoxy-2-[18F]fluoro-D-glucose myocardial uptake: effect of respiratory motion. Mol. Imaging Biol. 2003;5:57–64
5. Goerres GW, Kamel E, Heidelberg TN, Schwitter MR, Burger C and von Schulthess GK. PET-CT image co-registration in the thorax: influence of respiration. Eur. J. Nucl. Med. 2002;29:351–360
6. Goerres GW, Burger C, Kamel E, Seifert B, Kaim AH, Buck A, Buehler TC and von Schulthess GK. Respiration-induced attenuation artifact at PET/CT: technical considerations. Radiology 2003;226:906–910
7. Osman MM, Cohade C, Nakamoto Y, Marshall LT, Leal JP and Wahl RL. Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. J. Nucl. Med. 2003;44:240–243
8. Maintz JB and Viergever MA. A survey of medical image registration. Med. Image Anal. 1998;2:1–36
9. Fitzpatrick JM, West JB and Maurer CR. Predicting error in rigid-body point-based registration. IEEE Trans. Med. Imaging 1998;17:694–702

10. Wells WM, Viola P, Atsumi H and Nakajima S. Multi-modal volume registration by maximization of mutual information. Med. Image Anal. 1996;1:35–51

11. Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P and Marchal G. Automated multimodality image registration using information theory. Proc. Information Processing in Medical Imaging (Ile de Berder) (Dordrecht: Kluwer) 1995;263–274

12. Viola P and Wells WM. Alignment by maximization of mutual information. Proc. 5th Int. Conf. on Computer Vision (Boston) (Cambridge, MA: MIT Press) 1995;16–23

13. Maes F, Collignon A, Vandermeulen D, Marchal G and Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging 1997;16:187–198

14. Viola P and Wells WM. Alignment by maximization of mutual information. Int. J. Comput. Vision 1997;24:137–154

15. Duda R and Hart P. Pattern Classification and Scene Analysis (New York: Wiley). 1973

16. Ibanez L, Schroeder W, Ng L and Cates J. The ITK Software Guide: The Insight Segmentation and Registration Toolkit version 1.4 (New York: Kitware, Inc.). 2003

17. Schneider U, Pedroni E and Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. Phys. Med. Biol. 1996;41:111–124

18. Erdi YE et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). Radiother. Oncol. 2002;62:51–60

19. Klabbers BM, de Munck JC, Slotman BJ, Langendijk HA, de Bree R, Hoekstra OS, Boellaard R and Lammertsma AA. Matching PET and CT scans of the head and neck area: development of method and validation Med. Phys. 2002;29:2230–2238

20. Skalski J, Wahl RL and Meyer CR. Comparison of mutual information-based warping accuracy for fusing body CT and PET by 2 methods: CT mapped onto PET emission scan versus CT mapped onto PET transmission scan. J. Nucl. Med. 2002;43:1184–1187

21. Israel O, Mor M, Guralnik L, Gaitini D, Zachs Y, Keidar Z and Kuten A. The new technology of combined transmission and emission F-18 FDG tomography (FDG-PET) in the diagnosis and management of cancer patients. Clin. Positron Imaging 2000;3:143

22. Slomka PJ, Dey D, Przetak C, Aladl UE and Baum RP. Automated 3-dimensional registration of stand-alone (18)F-FDG whole-body PET with CT. J. Nucl. Med. 2003;44:1156–1167

23. Fitzpatrick JM and West JB. The distribution of target error in rigid-body, point-based registration. IEEE Trans. Med. Imaging 2001;20:917–927

24. Fitzpatrick JM, Hill DL, Shyr Y, West J, Studholme C and Maurer CR. Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain. IEEE Trans. Med. Imaging 1998;17:571–585

25. Rohlfing T and Maurer CR. Modeling liver motion and deformation during the respiratory cycle using intensitybased nonrigid registration of gated MR images. Med. Phys. 2004;31:427–432

26. Townsend DW and Beyer T. A combined PET/CT scanner: the path to true image fusion Br. J. Radiol. 2002;75:24–30

11

Evaluation of image registration in PET/CT of the liver, and recommendations for optimized imaging

Wouter V. Vogel [1]
Jorn A. van Dalen [1]
Bas Wiering [2]
HenkJan Huisman [3]
Frans H.M. Corstens [1]
Theo J.M. Ruers [2]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1] , Surgery [2], and Radiology [3],
Radboud University Nijmegen Medical Centre, the Netherlands

Journal of Nuclear Medicine, 2007
Volume 48 (6), pages 910-919.

Abstract

Multimodality PET/CT imaging of the liver can be performed with an integrated (hybrid) PET/CT scanner, or with software fusion of dedicated PET and CT. Regardless the method, accurate anatomical correlation and good image quality of both modalities are important prerequisits. Registration accuracy is influenced by breathing motion differences between PET and CT, which may also have impact on (attenuation correction related) artefacts, especially in the upper abdomen. The impact of these issues was evaluated for both hybrid PET/CT and software fusion, focussed on imaging of the liver.

Methods: 30 patients underwent hybrid PET/CT, 20 of whom with CT during expiration breathhold (EB) and 10 with CT during free breathing (FB). 10 additional patients underwent software fusion of dedicated PET and dedicated expiration breathhold CT (SF). The image registration accuracy was evaluated at the location of liver borders on CT and uncorrected PET images, and at the location of liver lesions. Attenuation correction artefacts were evaluated by comparison of liver borders on uncorrected and attenuation corrected PET images. CT images were evaluated for the presence of breathing artefacts.

Results: In EB, 40% of patients had an absolute registration error of the diaphragm in the vertical direction >1 cm (range -16 – 44 mm), and 45% of lesions were mispositioned >1 cm. In 50% of cases, attenuation correction artefacts caused a deformation of the dome of the liver of >1 cm. Poor compliance to breathhold instructions caused CT artefacts in 55% of cases. In FB, 30% had registration errors >1 cm (range -4 – 16 mm) and PET artefacts were less extensive, but all CT images suffered from breathing artefacts. As SF allows independent alignment of PET and CT no registration errors or artefacts >1 cm of the diaphragm occurred, and attenuation correction artefacts are not an issue.

Conclusions: When imaging the liver with hybrid PET/CT, registration errors and artefacts related to breathing motion may occur, depending largely on the selected breathing protocol and the speed of the CT scanner. According to these findings, recommendations were formulated with regard to scanner requirements, breathing protocols, and reporting.

11.1 Introduction

Accurate imaging of liver lesions such as metastases is important, as it has an impact on clinical decision making when considering partial liver resections and local ablative therapy such as radiofrequency ablation (RFA) (1,2). These locoregional interventions rely on accurate information about the localization and extent of tumor sites (3,4).

The added value of functional imaging with ^{18}F-fluoro-deoxy-glucose positron emission tomography (FDG-PET) to conventional anatomical imaging (especially CT and, MRI) has been well recognized, especially when assessing previous therapeutic interventions (5,6). However, exact localization and demarcation of lesions on FDG-PET is limited by a relatively low spatial resolution and lack of anatomical reference. The obvious benefit of combining the capabilities of CT (anatomical reference) and FDG-PET (sensitive tumor detection) has led to the practice of correlation of images as obtained by PET and by CT (7-9).

Correlation can be performed with mere visual evaluation of images acquired by separate scanners, or with integrated images as provided by either an integrated (hybrid) PET/CT scanner or software image fusion of dedicated PET and CT (10). Regardless of the methodology, the anatomical correlation of both image sets must be accurate. This implies that the liver needs to be in the same anatomical position and shape during both CT and PET acquisitions. However, CT and PET are differently influenced by breathing motion. PET is obligatory acquired during free breathing, and consequently suffers from some blurring in the region of the diaphragm. CT acquisition must be adapted to match these images, by scanning during free breathing or timed unforced expiration (10). Neither approach fully eliminates the risk of registration errors between PET and CT (11,12). Furthermore, the free breathing approach will introduce breathing artefacts on CT images in hybrid PET/CT, as attenuation correction of PET images is based on the CT images.

In this study, we evaluate the accuracy of anatomical registration and the occurence of artefacts during different breathing protocols in hybrid PET/CT imaging of the liver, and compare it to software image fusion of separately acquired PET and CT.

11.2 Methods and materials

Combined PET/CT images were acquired with 3 different protocols. 20 consecutive patients suspected of metastases from colorectal cancer underwent hybrid PET/CT with low-dose CT during expiration breathhold (EB). 10 other consecutive patients, who were referred for various indications and who were unable to comply to breathing instructions for various reasons, underwent hybrid PET/CT with low-dose CT during free breathing (FB). 10 more consecutive patients with suspected metastases from colorectal carcinoma underwent software fusion of dedicated PET and dedicated breathhold diagnostic CT (SF).

Image acquisition

Hybrid PET/CT scans were acquired using a Siemens Biograph Duo (Siemens/CTI, Knoxville, Tennessee, USA). A low-dose CT scan for localization and attenuation correction purposes was acquired in the caudo-cranial direction from the thighs to the skull base. Scanning parameters included 40 mAs, 130 kV, slice collimation 5 mm, pitch 1,5, reconstructed to 3 mm slices for smooth coronal representation. CT scans were acquired during timed unforced expiration breathhold (EB) or during free breathing (FB). Timed expiration breathhold consisted of free breathing during the caudal part of the scan, a deep inspiration command at the level of the spina iliaca superior, immediately followed by a command to expire and breathhold; patients were allowed to resume free breathing at the level of the lung tops. The total expiration breathhold time was about 30 seconds. Free breathing was performed without specific patient instructions. No intravenous contrast was applied. For PET scanning, a 3D emission scan of the central body was acquired during free breathing, 60 minutes after intravenous injection of 250 MBq FDG. The acquisition time per bed position was 4 minutes for emission only. Uncorrected emission images as well as images with CT-based attenuation correction were reconstructed, both using 2 iterations, 8 subsets and a 5 mm 3D Gaussian filter.

Dedicated FDG-PET scans were acquired using a Siemens ECAT Exact 47 scanner (Siemens/CTI, Knoxville, Tennessee, USA). A 3D emission scan was acquired and reconstructed identical to PET from PET/CT. In addition, a 2D Germanium-68 based transmission scan was acquired for attenuation correction. The acquisition time per bed position was 5 minutes for emission and 3 minutes for the transmission.

Dedicated CT scans were acquired using a Siemens Somatom Volume Zoom (Siemens, Erlangen, Germany). Scans of the liver were acquired with 80 mAs, 130 kV, slice thickness 5 mm, during unforced expiration breathhold for the whole scan length. Intravenous contrast was applied; the portal phase images were selected for image fusion with PET.

Image registration

For hybrid PET/CT scanning, normal image registration quality assurance procedures were followed as described by the manufacturer. This involved alignment of the PET and CT gantries after maintenance, using a crossed lines phantom. No additional image registration optimization was performed after scanning. Software image registration was performed on a PC with in-house developed image viewing and registration software based on the visualization toolkit

VTK (13) and the insight segmentation and registration toolkit ITK (14). The software allows rigid-body image registration, i.e. based on 3 translation and 3 rotation parameters. Anatomical registration of PET images to CT was pursued using an implementation of the automatic mutual information algorithm, restricted to a 3-dimensional volume containing the liver. The procedure has been described in more detail in a previous publication (15).

Definitions

PET images were evaluated for image registration errors, attenuation correction artefacts, and the visual discernibility of these errors. Analysis was performed on CT images, uncorrected PET images (uPET), and attenuation corrected PET images (acPET).

- *Registration errors* were defined as the relative anatomical/positional mismatch of structures (either circumscript lesions or organ borders) as visible on uPET and CT images, expressed as a distance in mm.
- *Attenuation correction artefacts* were defined as contour changes of structures as visible on PET images before and after attenuation correction, i.e. a difference of liver border positions between uPET and acPET images, expressed as a distance in mm.
- *Visible errors* were defined as mismatch of visible structures (circumscript lesions or organ borders) between acPET and CT, representing the extent in which the combination of registration errors and attenuation correction errors can be recognised and appreciated in acPET, expressed as a distance in mm.

Where focal lesions were evaluated, distances are expressed as 3-dimensional vectors. This is not possible where liver borders are evaluated, because an unidirectional shift of a liver border may be complicated by an (unrecognisable) deformation or rotation that alters the location that represents the top. Therefore, when evaluating the liver borders, distances were expressed as 1-dimensional distances along the axis of the largest movement (e.g. the vertical direction for the diaphragmatic dome and the caudal tip of the liver, and the lateral direction for the right lateral liver border). For all measured registration errors and attenuation correction artefacts, a distance of > 10 mm was considered potentially clinically relevant.

Registration error of liver borders

The tangent points (top) of the curved shape of three liver borders were selected as landmarks: the diaphragmatic dome, the right liver border, and the caudal tip. Three-dimensional ellipsoids were mapped to the curved shapes of the liver borders; the locations of the tangent points were then derived mathematically. This procedure was performed on CT, uPET and acPET images independently, blinded from eachother. The process of ellipsoid mapping is illustrated in figure 11.1.

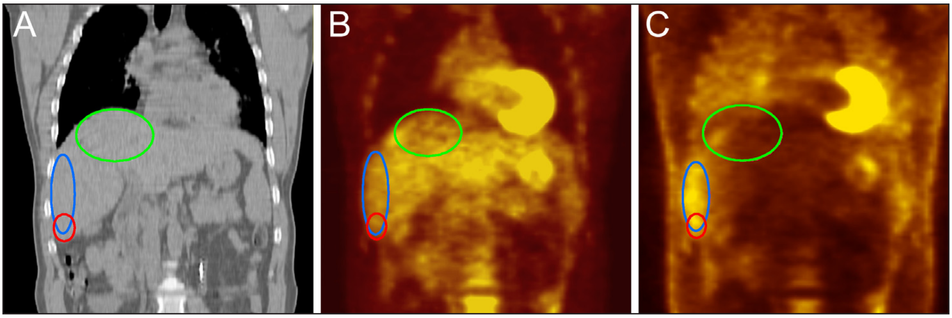


Figure 11.1
Localization of liver borders. Coronal slices of CT (A), attenuation corrected PET (B), and uncorrected PET (C) of a single patient, acquired with hybrid PET/CT during expiration breathhold. The circles represent slices through the 3-dimensional ellipsoids that were mapped to the diaphragmatic dome (green), right lateral border (blue) and the caudal tip (red), in order to determine differences in their respective positions.

An observer-specific systematical bias between localization of liver borders on uPET and acPET was determined by comparing measurements of images from dedicated PET, where the position of the liver is theoretically identical on both image sets. The true position of the liver border was assumed to be between the visual localizations on uPET and acPET. All uPET and acPET measurements were corrected afterwards for this bias, using the observer-specific average measurement difference from the theoretical position.

The interobserver variability for manual determination of positional differences of tangent points, after correction of the systematic bias, was evaluated in 5 subsequent dedicated PET scans (both uPET and acPET), by two experienced observers.

The error in local image registration was determined for the three liver borders separately, for all EB, FB, and SF images. Differences in image registration between imaging techniques were evaluated using Bartlett's test for equality of variances (level of significance 0.05).

Registration error of liver lesions

Image registration of focal lesions was evaluated using CT and acPET images, for EB and SF. Lesions were considered evaluable when they were visible on both CT and acPET. The center of each evaluable lesion was identified on CT and PET images. This analysis was not possible on free breathing CT images, as breathing motion artefacts on CT prevented reliable determination of the center of lesions. The interobserver variation in manual localization of lesion centers on CT and acPET was evaluated for 5 subsequent lesions on CT and acPET, by two experienced observers.

Attenuation correction artefacts

The extent of attenuation correction artefacts on PET was evaluated at the location of the liver borders. The apparent positions of the liver borders (tangent points) were determined before and after attenuation correction (i.e. on uPET and acPET images, respectively), similar to the evaluation of registration errors. This analysis was performed for EB, FB, and SF. Differences in the extent of attenation correction artefacts between protocols were evaluated using Bartlett's test for equality of variances (level of significance 0.05).

Visually discernible errors

The extent to which the combination of localization and attenuation correction errors were discernible on acPET was evaluated for all liver borders, by comparing acPET and CT in a similar approach as for the assessment of registration errors. This analysis was performed for EB, FB, and SF.

Breathing artefacts on CT

Artefacts caused by breathing motion may be depicted on CT images as locoregional deformities of the liver (i.e. breathhold not sustained) or as deformities throughout the liver (i.e. free breathing). The presence of both types of artefacts was evaluated visually for all CT images. Quantitative analysis of these artefacts was not attempted.

			Registration errors		**Attenuation correction artefacts**	
Breathing protocol	Measured landmark	Measured direction	Range (mm)	Absolute mean	Range (mm)	Absolute mean
EB	Diaphragmatic dome	Z-axis	-16 - 44	**11**	-18 - 41	**11**
	Right lateral border	X-axis	-8 - 8	**5**	-4 - 10	**2**
	Caudal tip	Z-axis	-3 - 53	**19**	-6 - 4	**2**
	Individual liver lesions	3D-vector	3 - 24	**11**		n.a.
FB	Diaphragmatic dome	Z-axis	-4 - 16	**7**	-7 - 11	**6**
	Right lateral border	X-axis	-4 - 7	**3**	-4 - 4	**2**
	Caudal tip	Z-axis	-5 - 20	**9**	-4 - 4	**2**
	Individual liver lesions	3D-vector		n.a.		n.a.
SF	Diaphragmatic dome	Z-axis	-3 - 8	**3**	-3 - 5	**3**
	Right lateral border	X-axis	-1 - 9	**3**	-2 - 2	**1**
	Caudal tip	Z-axis	-3 - 12	**5**	-3 - 4	**2**
	Individual liver lesions	3D-vector	7 - 14	**9**		n.a.

Table 11.1
The extent of registration errors and attenuation correction artefacts in PET/CT imaging of the liver, using different acquisition protocols.

11.3 Results

All FB scans, SF scans, and all registration procedures were performed without problems. From the 20 EB patients, 11 patients did not fully comply with adequate breathhold during CT acquisition of the whole liver range. This was visible on CT images as various artefacts, a more detailed evaluation is provided below. The EB group was therefore divided into 9 patients with adequate breathhold ($EB_{adequate}$) and 11 patients with inadequate breathhold (EB_{failed}) for additional separate analysis.

Registration error of liver borders

The absolute average interobserver variability in determination of liver border position differences on uPET and acPET, measured in one direction, was 2 mm (range -3 – 4 mm, S.D. 3 mm) for the diaphragmatic dome, 2 mm (range -3 – 2 mm, S.D. 2 mm) for the right lateral border, and 2 mm (range -4 – 3 mm, S.D. 3 mm) for the caudal tip.

For EB the absolute average image registration error at the diaphragmatic dome of the liver in the vertical direction was 11 mm (relative range -16 to +44 mm, 40% more than 10 mm). The largest errors were caused by insufficient expiration during breathhold CT. For the caudal tip of the liver, the average error was 19 mm (range 0 – 53 mm, 55% more than 10 mm). There were no clinically relevant registration errors at the right lateral liver border, and neither in the FB and SF protocols.

The image registration errors of FB and SF at the locations of the diaphragmatic dome and caudal tip were significantly less than that of EB ($P < 0.05$). SF did not perform significantly better than FB at the location of all liver borders. The results are listed in more detail in table 11.1. The distribution of registration errors per acquisition protocol is represented in figure 11.2. The occurrence of registration errors >10 mm is visualized in figure 11.3.

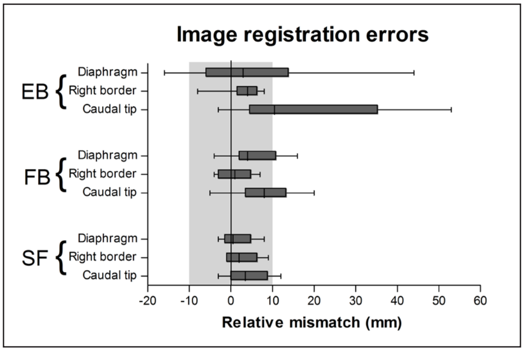


Figure 11.2
Image registration error of liver borders. The relative image registration errors at the location of several liver borders, for EB (hybrid PET/CT with breathhold CT), FB (hybrid PET/CT with free breathing CT), and for SF (software fusion of dedicated PET and CT). Registration errors occur mainly in the vertical direction (diaphragm and caudal tip affected most), due to insufficient expiration during CT.

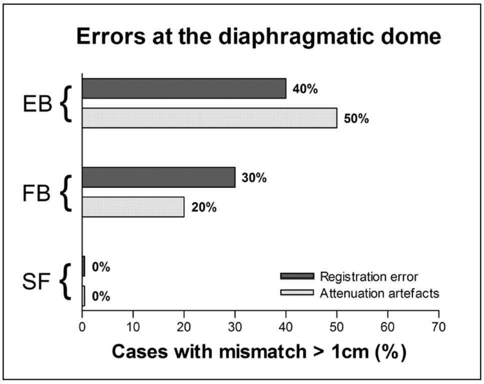


Figure 11.3
The frequency of image registration errors and attenuation correction artefacts of more than 10 mm at the location the diaphragmatic dome of the liver, for EB (hybrid PET/CT with breathhold CT), FB (hybrid PET/CT with free breathing CT), and for SF (software fusion of dedicated PET and CT).

The results of image registration at the diaphragmatic dome in breathhold PET/CT was not significantly influenced by the adequacy of the breathhold instructions (error > 10 mm in 44% of $EB_{adequate}$ and in 36% of EB_{failed}; not significant). Conversely, the registration of the caudal tip of the liver appeared to be influenced by the success of the breathhold procedure: error >10 mm in 33% of $EB_{adequate}$ and in 73% of EB_{failed}. Due to the small number of cases, this difference also failed to reach statical significance. When EB was limited to $EB_{adequate}$, there was no difference with FB in the diaphragmatic dome.

Registration error of liver lesions

The average interobserver variability in localization of focal liver lesions, measured as a 3-dimensional vector, was 2 mm (range 1 – 3 mm) on CT, and 1 mm (range 0 – 2 mm) on acPET. For EB, the average displacement of 11 detected lesions was 11 mm (range 3 – 24 mm), with 5 lesions (45%) being displaced more than 10 mm. There were insufficient evaluable lesions for separate analysis of $EB_{adequate}$ and EB_{failed}. For SF, the average displacement of 5 detected lesions was 9 mm (range 7 – 14 mm), with 1 lesion displaced more than 10 mm. Due to the limited number of evaluable lesions, statistical comparison of EB and SF was not performed. An example of a displaced lesion on hybrid PET/CT during EB is shown in figure 11.4.

Attenuation correction artefacts

For EB, the average size of attenuation artefacts at the diaphragmatic liver dome in the vertical direction was 11 mm (range 0 – 41 mm, 50% more than 10 mm). The largest errors were encountered due to insufficient expiration during breathhold CT. The extent of attenuation correction artefacts at the diaphragmatic dome did not depend on the success of the breathhold procedure (error > 10 mm in 45% of $EB_{adequate}$ and in 56% of EB_{failed}, not significant). An example of liver deformation due to attenuation correction in breathhold hybrid PET/CT is shown in figure 11.5.

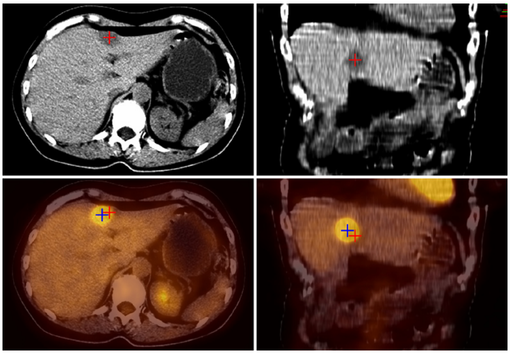


Figure 11.4
Misregistration of a liver lesion in breathhold PET/CT. Transverse (left) and coronal (right) images of a large liver metastasis in hybrid PET/CT with breathhold CT. The center of the lesion is marked with a red cross on CT, and with a blue cross on PET. Positioning differences of the liver between PET and CT acquisition resulted in a mismatch of 13 mm, measured as a 3-dimensional vector.

For FB, the average attenuation artefact at the diaphragmatic liver dome measured 6 mm (range 0 – 11 mm, 20% more than 10mm), both in in the cranial and caudal direction. For SF, attenuation correction artefacts of the liver are theoretically not an issue as attenuation is performed with Germanium rod sources during free breathing. Control measurements at the diaphragmatic dome showed an average absolute error of 3 mm (range of 0 – 8 mm, thus in all patients within 10 mm). No signifant attenuation correction artefacts occurred at the lateral border or the caudal tip, for all EB, FB and SF cases.

The extent of attenuation correction artefacts at the diaphragmatic dome was significantly worse in EB than in FB or SF ($P < 0.05$). When EB was limited to $EB_{adequate}$, the difference with FB in the diaphragmatic dome was no longer significant. FB was significantly worse than SF ($P < 0.05$). No clinically relevant attenuation correction errors occurred at the right lateral border and the caudal tip with either technique.

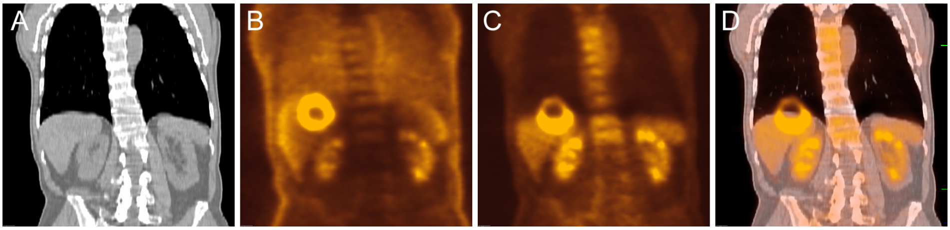


Figure 11.5
PET/CT attenuation correction artefacts. Hybrid PET/CT of a large liver metastasis with central necrosis, with CT acquired during expiration breathhold. Coronal slices of CT (A), uncorrected PET (B), attenuation corrected PET (C), and fused corrected PET with CT (D). Despite breathing instructions, comparison of A and B reveals a difference in diaphragm position between CT and PET. C and D demonstrate a change in shape of the liver on PET after attenuation correction, to falsely match CT. The liver metastasis appears partially in the lung on corrected images, and suffers from severe loss of signal intensity in the region of the mismatch.

Visually recognisable errors

At the locations of the diaphragm and the right lateral liver border, there were no cases that showed a visually discernible mismatch of more than 10 mm at the liver border between acPET and CT, for all EB, FB and SF images, regardless of the presence of registration or attenuation correction artefacts of more than 10 mm. Visually discernible errors were seen at the the location of the caudal tip in all image series, with values similar to the local image registration error.

CT artefacts

In EB, locoregional breathing artefacts in the liver were detected in the CT images of 4 patients (20%), all attributable to some breathing motion during acquisition despite instructions to hold the breath. Furthermore, the caudal tip of the right liver lobe appeared displaced or deformed in 7 additional patients (35%), all attributable to the breathing instructions given when approaching the region of the liver in the caudocranial scanning direction. In FB, free breathing artefacts were discernible throughout the images for all patients. In the SF images, no breathing artefacts were detected on CT images. Examples of breathing artefacts on CT are shown in figure 11.6. The clinically relevant breathing artefacts on CT are summarized in table 11.2.

Figure 11.6
Breathing artefacts on CT. CT slices from different patients, acquired on a hybrid PET/CT scanner. (A) Coronal slice of CT acquired with expiration breathhold command. The arrows indicate an artefact in the middle of the liver and spleen due to unsustained breathhold. (B) Sagittal slice of CT acquired with expiration breathhold command. The breathing commands were given relatively late, and can be recognised by movement of the abdominal wall (left arrow); the resulting liver motion causes the caudal tip of the liver to appear twice (right arrow). (C) Coronal slice of CT acquired during free breathing. Breathing artefacts are visible throughout the image.

11.4 Discussion

Evaluation of PET/CT image registration and artefacts of the liver is not trivial, because the organ lacks well-defined, clearly discernible landmarks on PET. Therefore, evaluations are limited to liver borders, and if present, focal lesions within the liver. Evaluation of liver borders is further restricted to those areas with sufficient contrast to surrounding tissues in uncorrected PET images. acPET images from PET-CT can not be used for this purpose because CT-based attenuation correction may influence the images. Therefore, the top of the diaphragmatic dome, the lateral edge of the right liver border, and the caudal tip of the right liver lobe were selected for evaluation. Other borders of the liver could not be evaluated using uPET.

Comparison of border localizations is complicated as liver borders appear different on uPET and acPET. For example, on acPET the level of the diaphragm shows a sharp transition from low to high uptake (lungs to liver), while uPET shows a transition from medium intensity in the lungs to depth-dependent variable intensity in the liver. As this may have caused a systematic bias in localization of liver borders, an observer-specific correction factor needed to be applied

	Hybrid PET/CT Breathhold CT		**Hybrid PET/CT Free breathing CT**		**Software image fusion**	
Artefacts on CT	**Affected cases**		**Affected cases**		**Affected cases**	
- Breathhold not sustained	3 / 20	**20 %**	0 / 10	**---**	0 / 10	**0 %**
- Breathhold timing issues	5 / 20	**33 %**	0 / 10	**---**	0 / 10	**0 %**
- Free breathing artefacts	0 / 20	**---**	10 / 10	**100 %**	0 / 10	**---**
Image registration errors > 1 cm						
- Diaphragmatic dome	8 / 20	**40 %**	3 / 10	**30 %**	0 / 10	**0 %**
- Right lateral border	0 / 20	**0 %**	0 / 10	**0 %**	0 / 10	**0 %**
- Caudal tip	11 / 20	**55 %**	4 / 10	**40 %**	2 / 10	**20 %**
- Individual liver lesions	5 / 11	**45 %**	---	---	1 / 5	**20 %**
Attenuation correction artefacts on PET > 1 cm						
- Diaphragmatic dome	10 / 20	**50 %**	2 / 10	**20 %**	0 / 10	**0 %**
- Right lateral border	0 / 20	**0 %**	0 / 10	**0 %**	0 / 10	**0 %**
- Caudal tip	0 / 20	**0 %**	0 / 10	**0 %**	0 / 10	**0 %**

Table 11.2
Errors and artefacts of more than 10 mm in multimodality PET/CT imaging of the liver.

to all measurements of uPET and acPET. The correction factor was derived from a series of measurements on uPET and acPET from dedicated PET (where the position of the liver should be identical); the true position of a liver border was assumed to be at the mathematical middle of measurements. The corrected measurements may still differ slightly from the real border positions, but the differences will be small and cannot be optimized any further.

Despite correction of any systematical bias, the manual localization of liver borders and focal lesions can never be perfect. Uncertainties are caused by the limited spatial and contrast resolution of PET, and by interpretation difficulties on uPET images in general. Mapping of 3-dimensional ellipsoids to liver borders may reduce sampling errors to some extent, but cannot fully eliminate them. Furthermore, different observers may choose different points of the liver for the top, due to the sometimes irregular shape of the organ. Therefore, the interobserver variability measurements were restricted to comparison of positional differences rather than positions of liver borders, thus eliminating the variable choice of the top as a factor. The measured interobserver variabilities were all well within acceptable ranges (2 mm on average between uPET and acPET, for all borders).

Obviously, measurement of the available landmarks (3 borders in one direction each and a limited number of focal lesions) represents an oversimplification of the true position of the liver. Only basic liver displacement will be detected, while local deformation and organ rotation will be neglected. Thus, the observed misregistrations in this study rather represent an underestimation than an overestimation.

The detection limit of FDG-PET for small liver lesions has been estimated in the range of 10 mm (16). Thus, misregistration needs to be more than ~10 mm to cause uncertainty in discrimiation of two adjacent small structures. This does not imply that all cases with registration errors >10 mm will lead to misinterpretations, but awareness of the (possible) extent of misregistration may help to avoid reading errors.

Breathing issues

The problem of maintaining unforced expiration breathhold is easily underestimated. The procedure is demanding and needs to be rehearsed prior to scanning. Even so, some patients will fail to sustain breathhold during actual scanning, causing CT artefacts (in 20% of cases in our series). Especially when scanning elderly or diseased patients, completion of a CT scan during breathhold cannot be guaranteed. This problem is obviously related to the the acquisition time for CT during whole-body scanning. The faster the CT scanner, the better breathhold compliance will be.

Another cause of artefacts on breathhold CT images is the timing of breathing instructions: 33% of cases showed deformation of the caudal part of the liver, related to the deep inspiration command when approaching the mid-abdomen. Depending on the speed of the CT scanner, an earlier breathhold command would result in increased risk on breathing artefacts towards the end of the scan in the upper lung fields. Increasing the CT acquisition speed may further degrade the quality of the low-dose CT images.

The alternative strategy of free superficial breathing resulted in CT artefacts throughout the liver. This was previously described by Beyer et al. (11). In the lungs, this effect caused small lung nodules to be missed in up to 34% of cases (17). It is unknown how this translates to imaging of the liver, but such a level of missed diagnosis will be hardly acceptable for correlative imaging. A faster CT scanner will result in free breathing artefacts with a lower frequency, but the amplitude of misregistrations will remain unchanged as this depends on patient factors only. Thus, both breathhold and free breathing techniques have disadvantages.

Even in an ideal situation (i.e. with a fast scanner, accurate breathhold instructions, and an exemplary patient), the exact position of the diaphragm during breathhold cannot be predicted. Furthermore, the shape of the diaphragm may vary because breathhold generates a different muscle tension than free breathing. This implies that differences in position and shape of the liver between PET and CT may be unavoidable to some extent. Our results confirm that registration errors of the liver are not uncommon, and are rather unpredictable in extent. The misregistrations occurred mainly in the cranio-caudal direction, and were in most cases explained by insufficient expiration during breathhold. Deeper expiration would increase the risk of non-sustained breathhold. Perfect compliance to breathhold instructions did not prevent the occurrence of registration errors >10 mm, although the average extent appeared lower. The impact of breathhold protocols has been evaluated by Goerres et al. (18). They concluded that the best breathing protocol is unforced expiration breathhold as performed in our study, but that the impact on image registration can still be severe (relative registration errors of −25 to +19 mm; compare −16 to +44 mm in our series). Brechtel et al. have reported better values for image registration at the diaphragm, but these data seem biased because evaluation was limited to acPET images only (19).

Free breathing during CT resulted in less extreme registration errors, but still >10 mm in 30%. The extent of misregistration was congruent with normal liver motion during free breathing as shown by Brandner et al. (20). As stated above, these registration errors will be unavoidable. The extent of misregistration is theoretically not dependent on the speed of the CT scanner, but this could not be evaluated in this study. The effects of free breathing CT on image registration of abdominal organs in PET/CT have previously been evaluated by Nakamoto et al. (21). They observed even slightly worse results at the location of the right diaphragmatic dome, with

38% misregistration >10 mm and even 10% of cases >20 mm (30% and 0% respectively in our study), with artefacts that may influence the position, shape and apparent size of the liver on PET. Osman et al. demonstrated moderate to severe attenuation correction artefacts in 18% of cases at the right diaphragmatic dome, although quantitative analysis was not performed (22). They also observed that correlation errors of liver lesions may occur, incidentally even with erroneous localization in lung instead of liver, albeit in a limited number of cases (23). Papathanassiou et al. have confirmed that lesions may be missed in liver parts that were affected by CT-based attenuation correction (24).

Selecting a protocol

Given the possible extent of lesion misregistration and attenuation correction artefacts, integrated PET/CT can be far from ideal for imaging of the liver. Free breathing hybrid PET/CT performs somewhat better at these issues, but suffers from poorer image correlation due to artefacts and missing of lesions on CT. Nevertheless, both approaches have been found suitable for diagnostic correlative imaging (25). When considering a breathhold or a free breathing hybrid PET/CT protocol, it is important realize that registration errors and attenuation correction artefacts in breathhold PET/CT can be recognised and circumvented afterwards by adequate evaluation of uPET images, while missing small lesions on free breathing CT is definitive. The final choice of technique may be guided by specific clinical questions, available equipment, individual patient characteristics, and personal preferences.

Software fusion is based on optimization of image registration of diagnostic quality image sets. This technique resulted in significantly lower registration errors of the liver as a whole, and of liver lesions, as PET and CT images can be freely manipulated to achieve optimal matching. This raises the question whether such optimization and uncompromised PET image quality are possible in hybrid PET/CT. When image registration is performed with uPET and CT images, followed by the attenuation correction procedure, the result should be similar to SF. However, this option is not provided by current hybrid PET/CT scanners. Software fusion is not ideal for high-throughput imaging, but remains useful when hybrid PET/CT is not available.

Eliminating attenuation correction artefacts

When the abovementioned registration mismatches between PET and CT occur, hybrid PET/CT introduces a subsequent risk for artefacts on PET images. Attenuation correction will be applied erroneously on PET at the location of dense objects (26,27), or where the location of a transition from low to high photon attenuating tissue does not correspond on PET and CT (21,28). The diaphragmatic area is very susceptible to such errors due to the sharp tissue/air transition, combined with the risk for positional differences. This may result in an apparent contour change of the liver on acPET images, and reduced sensitivity for lesions in the affected area.

In our series of EB imaging, deformation of the liver on acPET images was not uncommon, and rather unpredictable. The potential clinical impact is underlined by the presence of artefacts >10 mm in 50% of cases. In FB attenuation correction artefacts were significantly less extensive, but still in 20% >10 mm. These problems must be considered unavoidable, as long as registration errors occur and attenuation correction is performed with CT images. Dedicated PET does not easily suffer from attenuation correction artefacts, because the attenuation profile is measured using photons with an energy identical to emission scanning,

Category	**Recommendations**
Scanner requirements	When hybrid PET/CT with breathhold CT is needed, a fast CT scanner (i.e. more than 2 slices) is preferable to avoid breathhold compliance issues. When hybrid PET/CT with a fast CT is available, breathhold CT seems preferable over free breathing CT (no missed lesions and no artefacts on CT, while image registration and artefacts are not significantly worse than in free breathing). When hybrid PET/CT with a slow CT is available, the choice between breathhold and free breathing CT (for non-diagnostic use) is unsettled and depends on personal preference (i.e. more serious registration errors with breathhold CT, but missed lesions on free breathing CT).
PET/CT image acquisition	Explain the importance of avoiding motion to patients Practising breathhold with each patient prior to actual scanning is advised, to avoid serious misregistration and artefacts. Revert to free breathing when breathhold rehearsal fails. The performance of PET/CT with breathhold CT may be improved by providing feedback about registration errors to operating personnel. When reliable correlative imaging of PET and CT images without artefacts is needed on an incidental basis, software fusion of dedicated PET and diagnostic CT can still be considered.
Reviewing PET/CT	Awareness of the level of misregistration and attenuation correction artefacts can be improved by consequent correlation of uncorrected PET and CT images. Uncorrected PET images may reveal small lesions that may be undetectable or misplaced on corrected PET images, in the diaphragmatic area of the liver and the lower lung fields. Unexplained PET lesions that show no correlating density on non-diagnostic CT (e.g. free breathing or low-dose) may be resolved by correlation with separate diagnostic CT images.

Table 11.3
Recommendations for optimal PET/CT imaging and reviewing. These recommendations were formulated to achieve optimal diagnostic quality and interpretation of multimodality PET/CT imaging of the liver. As such, most recommendations will also apply to whole-body imaging with PET/CT.

acquired during the same breathing pattern. No significant artefacts were detected in our series of dedicated PET images. This leaves room for improvement of hybrid PET/CT with regard to PET attenuation correction artefacts. As full elimination of registration errors seems unlikely, re-introduction of 511 KeV transmission imaging in hybrid PET/CT remains a matter of debate. Expiration breathhold on CT images will remain a problem for correlative imaging. Alternative acquisition protocols all bear disadvantages. Slow-CT and averaged multiple-series-CT as applied in external beam radiation therapy planning severely degrade the image quality. Gated CT acquisition may provide excellent correlative imaging, at least when PET is acquired in gating identical to CT imaging (29). Further experiments with such techniques need to be conducted.

Future developments

Recent advances in CT scanning have predominantly been in the field of spatial resolution and scanning speed, with an increasing number of slices (e.g. 128 parallel detectors, multibeam CT). A real benefit could be expected from better attenuation correction. Although algorithms for CT-based attenuation correction are continuously being improved, only reintroduction of true 511 KeV transmission imaging in hybrid PET/CT scanners would minimize the problem of artefacts. This remains a matter of debate, although some hybrid PET/CT scanners offer this functionality. Development of higher quality and faster transmission scanning, e.g. simultaneous with emission scanning (30), is eagerly awaited.

Other PET radiopharmaceuticals for imaging of malignancy in the liver are likely to become available in the coming years, such as ^{18}F-fluor-deoxy-thymidine (FLT) (31). Registration errors and attenuation correction artefacts in FDG-PET/CT imaging are independent of the PET tracer. However, misregistrations will be harder to detect in tracers that show low or no uptake in normal liver tissue. This illustrates the importance of optimized imaging and reviewing, especially for imaging of novel tracers in the near future.

The combination of PET with MRI may be preferable over PET/CT, for better soft tissue imaging characteristics and less radiation dose issues, but breathing will remain a problem. Current MRI techniques do not allow whole-body imaging during breathhold, and free breathing during MRI can severely distort the images. The best approach to hybrid PET/MRI is still unclear.

11.5 Conclusions

Anatomical registration errors of the liver in PET/CT may be severe, occuring mainly in the vertical direction, due to breathing differences during acquisition of PET and CT. Subsequent attenuation correction artefacts can occur where registration errors are present, but only where a sharp transition between dense and non-dense tissue exists (e.g. at the diaphragmatic area but not in the caudal region of the liver). The extent of image registration errors and attenuation correction errors cannot be appreciated visually on PET images that have been corrected for photon attenuation.

The application of breathing protocols for CT images influences these issues, but no protocol can warrant perfect image registration and artefact-free images. PET/CT with breathhold CT may suffer from unpredictable significant misregistrations. A fast CT scanner is less liable to breathhold compliance errors and timing issues and may therefore reduce serious registration errors. PET/CT with free breathing CT also suffers registration errors, and from unavoidable breathing artefacts throughout CT images which must therefore be considered non-diagnostic.

Awareness of the level of misregistration and attenuation correction artefacts is essential for reviewing, and can be improved by consequent correlation of uncorrected PET and CT images. Furthermore, uncorrected PET images may allow detection of small lesions that became invisible or misplaced on corrected PET images, in the diaphragmatic area of the liver and the lower lung fields. Based on these conclusions, recommendations were formulated for optimal imaging and reviewing of integrated PET/CT (table 11.3).

11.6 References

1. Buscarini E, Savoia A, Brambilla G, Menozzi F, Reduzzi L, Strobel D, Hansler J, Buscarini L, Gaiti L, and Zambelli A. Radiofrequency thermal ablation of liver tumors. Eur. Radiol. 2005;15:884-894
2. Gayowski TJ, Iwatsuki S, Madariaga JR, Selby R, Todo S, Irish W, and Starzl TE. Experience in hepatic resection for metastatic colorectal cancer: analysis of clinical and pathologic risk factors. Surgery. 1994;116:703-710
3. Blokhuis TJ, van der Schaaf MC, van den Tol MP, Comans EF, Manoliu RA, and van dS, Jr.. Results of radio frequency ablation of primary and secondary liver tumors: long-term follow-up with computed tomography and positron emission tomography-18F-deoxyfluoroglucose scanning. Scand. J. Gastroenterol. Suppl. 2004;93-97
4. Podnos YD, Henry G, Ortiz JA, Ji P, Cooke J, Cao S, and Imagawa DK. Laparoscopic ultrasound with radiofrequency ablation in cirrhotic patients with hepatocellular carcinoma: technique and technical considerations. Am. Surg. 2001;67:1181-1184
5. Antoch G, Vogt FM, Veit P, Freudenberg LS, Blechschmid N, Dirsch O, Bockisch A, Forsting M, Debatin JF, and Kuehl H. Assessment of liver tissue after radiofrequency ablation: findings with different imaging procedures. J. Nucl. Med. 2005;46:520-525
6. Dromain C, de BT, Elias D, Kuoch V, Ducreux M, Boige V, Petrow P, Roche A, and Sigal R. Hepatic tumors treated with percutaneous radio-frequency ablation: CT and MR imaging follow-up. Radiology. 2002;223:255-262

7. Kamel IR, Cohade C, Neyman E, Fishman EK, and Wahl RL. Incremental value of CT in PET/CT of patients with colorectal carcinoma. Abdom. Imaging. 2004;29:663-668

8. Selzner M, Hany TF, Wildbrett P, McCormack L, Kadry Z, and Clavien PA. Does the novel PET/CT imaging modality impact on the treatment of patients with metastatic colorectal cancer of the liver? Ann. Surg. 2004;240:1027-1034

9. Veit P, Antoch G, Stergar H, Bockisch A, Forsting M, and Kuehl H. Detection of residual tumor after radiofrequency ablation of liver metastasis with dual-modality PET/CT: initial results. Eur. Radiol. 2005;

10. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, and Corstens FH. PET/CT: panacea, redundancy, or something in between? J. Nucl. Med. 2004;45 Suppl 1:15S-24S

11. Beyer T, Antoch G, Blodgett T, Freudenberg LF, Akhurst T, and Mueller S. Dual-modality PET/CT imaging: the effect of respiratory motion on combined image quality in clinical oncology. Eur. J. Nucl. Med. Mol. Imaging. 2003;30:588-596

12. Kim JH, Czernin J, len-Auerbach MS, Halpern BS, Fueger BJ, Hecht JR, Ratib O, Phelps ME, and Weber WA. Comparison between 18F-FDG PET, in-line PET/CT, and software fusion for restaging of recurrent colorectal cancer. J. Nucl. Med. 2005;46:587-595

13. Schroeder W. The Visualization Toolkit: An object-oriented approach to 3D graphics, 3rd edn. New York: Kitware Inc. 2003;

14. Ibanez L, Schroeder W, Ng L, and Cates J. The ITK Software Guide: The Insight Segmentation and Registration Toolkit, version 1.4. New York: Kitware, Inc. 2003;

15. van Dalen JA, Vogel W, Huisman HJ, Oyen WJ, Jager GJ, and Karssemeijer N. Accuracy of rigid CT-FDG-PET image registration of the liver. Phys. Med. Biol. 2004;49:5393-5405

16. Fong Y, Saldinger PF, Akhurst T, Macapinlac H, Yeung H, Finn RD, Cohen A, Kemeny N, Blumgart LH, and Larson SM. Utility of 18F-FDG positron emission tomography scanning on selection of patients for resection of hepatic colorectal metastases. Am. J. Surg. 1999;178:282-287

17. Allen-Auerbach M, Yeom K, Park J, Phelps M, and Czernin J. Standard PET/CT of the Chest During Shallow Breathing Is Inadequate for Comprehensive Staging of Lung Cancer. J. Nucl. Med. 2006;47:298-301

18. Goerres GW, Kamel E, Heidelberg TN, Schwitter MR, Burger C, and Von Schulthess GK. PET-CT image co-registration in the thorax: influence of respiration. Eur. J. Nucl. Med. Mol. Imaging. 2002;29:351-360

19. Brechtel K, Klein M, Vogel M, Mueller M, Aschoff P, Beyer T, Eschmann SM, Bares R, Claussen CD, and Pfannenberg AC. Optimized contrast-enhanced CT protocols for diagnostic whole-body 18F-FDG PET/CT: technical aspects of single-phase versus multiphase CT imaging. J. Nucl. Med. 2006;47:470-476

20. Brandner ED, Wu A, Chen H, Heron D, Kalnicki S, Komanduri K, Gerszten K, Burton S, Ahmed I, and Shou Z. Abdominal organ motion measured using 4D CT. Int. J. Radiat. Oncol. Biol. Phys. 2006;65:554-560

21. Nakamoto Y, Tatsumi M, Cohade C, Osman M, Marshall LT, and Wahl RL. Accuracy of image fusion of normal upper abdominal organs visualized with PET/CT. Eur. J. Nucl. Med. Mol. Imaging. 2003;30:597-602

22. Osman MM, Cohade C, Nakamoto Y, and Wahl RL. Respiratory motion artifacts on PET emission images obtained using CT attenuation correction on PET-CT. Eur. J. Nucl. Med. Mol. Imaging. 2003;30:603-606

23. Osman MM, Cohade C, Nakamoto Y, Marshall LT, Leal JP, and Wahl RL. Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. J. Nucl. Med. 2003;44:240-243

24. Papathanassiou D, Becker S, Amir R, Meneroux B, and Liehn JC. Respiratory motion artefact in the liver dome on FDG PET/CT: comparison of attenuation correction with CT and a caesium external source. Eur. J. Nucl. Med. Mol. Imaging. 2005;32:1422-1428

25. Goerres GW, Burger C, Schwitter MR, Heidelberg TN, Seifert B, and Von Schulthess GK. PET/CT of the abdomen: optimizing the patient breathing pattern. Eur. Radiol. 2003;13:734-739

26. Dizendorf E, Hany TF, Buck A, Von Schulthess GK, and Burger C. Cause and magnitude of the error induced by oral CT contrast agent in CT-based attenuation correction of PET emission studies. J. Nucl. Med. 2003;44:732-738

27. Goerres GW, Ziegler SI, Burger C, Berthold T, Von Schulthess GK, and Buck A. Artifacts at PET and PET/CT caused by metallic hip prosthetic material. Radiology. 2003;226:577-584

28. Goerres GW, Burger C, Kamel E, Seifert B, Kaim AH, Buck A, Buehler TC, and Von Schulthess GK. Respiration-induced attenuation artifact at PET/CT: technical considerations. Radiology. 2003;226:906-910

29. Wolthaus JW, van HM, Muller SH, Belderbos JS, Lebesque JV, de Bois JA, Rossi MM, and Damen EM. Fusion of respiration-correlated PET and CT scans: correlated lung tumour motion in anatomical and functional scans. Phys. Med. Biol. 2005;50:1569-1583

30. Matsumoto K, Kitamura K, Mizuta T, Tanaka K, Yamamoto S, Sakamoto S, Nakamoto Y, Amano M, Murase K, and Senda M. Performance Characteristics of a New 3-Dimensional Continuous-Emission and Spiral-Transmission High-Sensitivity and High-Resolution PET Camera Evaluated with the NEMA NU 2-2001 Standard. J. Nucl. Med. 2006;47:83-90

31. Francis DL, Freeman A, Visvikis D, Costa DC, Luthra SK, Novelli M, Taylor I, and Ell PJ. In vivo imaging of cellular proliferation in colorectal cancer using positron emission tomography. Gut. 2003;52:1602-1606

12

A novel iterative method for lesion delineation and volumetric quantification with FDG-PET

Jorn A. van Dalen [1]
Aswin Hoffmann [2]
Volker Dicken [3]
Wouter V. Vogel [1]
Bastiaan Wiering [4]
Theo J. Ruers [4]
Nico Karssemeijer [5]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1] , Radiation oncology [2], Surgery [4], and Radiology [5],
Radboud University Nijmegen Medical Centre, the Netherlands.
MeVis-center for medical diagnostic systems and visualisation [3], Bremen, Germany.

Nuclear Medicine Communications, 2006
Volume 28 (6), pages 485-493.

12 Abstract

Determination of lesion boundaries on FDG-PET is difficult due to the point-spread blurring and unknown uptake of activity within a lesion. Standard threshold-based methods for volumetric quantification on PET usually neglect any size dependence and are biased by dependence on the signal-to-background ratio (SBR). A novel, model-based method is hypothesized to provide threshold levels independent of the SBR and to allow accurate measurement of volumes down to the resolution of the PET scanner.

Methods: A background-subtracted relative-threshold level (RTL) method was derived, based on a convolution of the point-spread function and a sphere with diameter D. Validation of the RTL-method was performed using PET imaging of a Jaszczak phantom with seven hollow spheres (D=10-60 mm). Activity concentrations for the background and spheres (signal) were varied to obtain SBRs of 1.5-10. An iterative procedure was introduced for volumetric quantification, as the optimal RTL depends on a priori knowledge of the volume. Feasibility of the RTL-method was tested in two patients with liver metastases and compared to a standard method using a fixed percentage of the signal.

Results: Phantom data validated that the theoretically optimal RTL depends on the sphere-size, but not on the SBR. Typically, RTL=40% (D=15-60 mm), and RTL>50% for small spheres (D<12 mm). The RTL-method is better applicable to patient data than the standard method.

Conclusions: Based on an iterative procedure, the RTL-method has been shown to provide optimal threshold levels independent of the SBR and to be applicable in phantom and in patient studies. It is a promising tool for lesion delineation and volumetric quantification of PET lesions.

12.1 Introduction

Metabolic information provided by positron emission tomography with ^{18}F-fluor-deoxy-glucose (FDG-PET) is increasingly used in the management of patients with cancer. Especially the role of FDG-PET in target volume delineation in radiotherapy treatment planning, both for the primary tumor and regional lymph nodes, is a main topic of research (1). Volumetric quantification with PET may also impact on surgical decision making (2).

When evaluating liver metastases, accurate imaging of such lesions is important for determining the feasibility of partial liver resections and especially of local ablative therapy such as radiofrequency ablation (RFA) (3,4). These locoregional interventions rely on accurate information about the localization and extent of tumor sites (5,6). The added value of FDG-PET to conventional anatomical imaging such as computed tomography (CT) and magnetic resonance imaging (MRI) for assessment of patients with liver metastases has been well documented (7,8).

FDG-PET is suitable to identify a tumor and its probable location. However, due to the point-spread blurring and the unknown amount of accumulated activity within the lesion, it is difficult to quantify its boundary and volume. The delineation of the "true" contour of the FDG-accumulation is addressed in different ways by various groups, such as visual interpretation (9,10), the application of a contour comprising a percentage of the maximum activity concentration in the lesion (9,11,12), or the use of a threshold based on the standardized uptake value (SUV) of e.g. 2.5 (9,13). This is what is mostly done in normal clinical practice. Some propose a more sophisticated method where the ratio of activity concentration in the lesion (signal S) to the concentration in normal tissue (background B) is the main input parameter in the contour detection (14,15). Others propose a method where the difference "S-B" is the main variable (16-18).

In this study, we developed a robust and relatively simple model-based method for lesion delineation and volumetric quantification with PET, where the optimal threshold levels take into account the presence of background activity and are independent of the signal-to-background ratio (SBR). The background of our method was based on a model for volumetric quantification with single-photon emission computed tomography (SPECT) (19,20). However, a translation to a methodology for data from FDG-PET, taking also into account the presence of background activity, has not been described before. Furthermore, an explicit derivation of the optimal threshold level that is independent of the SBR has not been provided before. Also, no procedure has been described how to implement such methodology in practice.

12.2 Theoretical background

In a one-dimensional situation, figure 12.1 illustrates how a step in an ideal (i.e. homogeneous) intensity distribution is displayed in an image. Due to the finite spatial resolution of the system that acquires the image, the edges of the step are blurred in the image. To obtain the true size of the step (D), it is easy to see that the 50%-level of the image signal (S) after background (B) subtraction indicates the optimal image cut-off value for the threshold (T). The only criteria are that D is sufficiently larger than the full-width-half-maximum value (FWHM) of the point-spread function (PSF) and that the PSF is symmetrical.

In a three-dimensional (3D) situation, mathematics become more complex since curved boundaries are encountered. In case of a sphere with diameter D, the optimal relative threshold level (RTL) to obtain the true volume, defined as the percentage of "S-B" where the cut-off should be taken, turns out to be smaller than 50%, for D approximately 1.2 times larger than the FWHM of the PSF (20). This can be seen by convolving a sphere with a trivariate Gaussian PSF, to determine the optimal RTL. The basic idea of the convolution is to determine the amount of overlap (an integral) between the PSF and the sphere, as a function of the distance between their centres (19,20). Firstly, the activity concentration, A_c, after background subtraction can be written as:

$$A_c(x_p,S,FWHM) = \int\int\int O\,(D) \times PSF\,(x_p,FWHM)\, dxdydz \qquad (12.1)$$

where $PSF\,(x_p,FWHM)$ is the point spread function centered at location x_p along the x-axis, with a full-width-half-maximum value FWHM, and $O\,(D)$ is the sphere with diameter D, centered at $x = y = z = 0$. Secondly, to determine the optimal RTL one should realize that the maximum amount of overlap between the sphere and the PSF is obtained when they are co-centered for $x_p = 0$. This implies that the image signal after background subtraction can be written as:

$$S - B = A_c(x_p = 0,D,FWHM) \qquad (12.2)$$

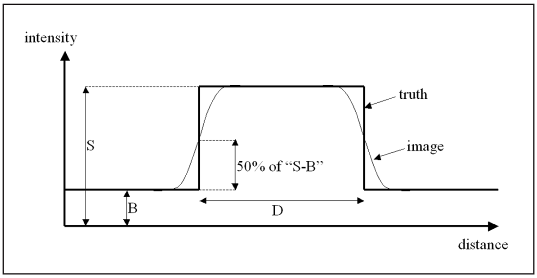


Figure 12.1
Illustration of a one-dimensional step-distribution, that is blurred in the image due to the point-spread function (PSF). The 50%-level of the difference "S-B" indicates the optimal image cut-off value (relative threshold level, RTL) to obtain the true size of the step (D).

In a similar way, the optimal threshold, after background subtraction, is obtained when x_p is equal to the radius of the sphere, i.e.:

$$T - B = A_c(x_p = ½D, D, FWHM) \tag{12.3}$$

Subsequently, the optimal RTL is obtained via:

$$\begin{aligned} RTL &= (T-B) / (S-B) \\ &= A_c(x_p = ½D, D, FWHM) / A_c(x_p = 0, D, FWHM) \end{aligned} \tag{12.4}$$

Figure 12.2 illustrates the optimal RTL as a function of the diameter of the sphere. In the calculations a trivariate Gaussian PSF with a FWHM of 9.2 mm (in all 3 dimensions) was used as input (for reasons explained below). This theoretical derivation forms the basis of this study.

The theoretical derivation shows that the optimal RTL is independent of the SBR. However, a dependency to the volume (or diameter) of the sphere is expected. As shown in figure 12.2, for very large spheres, RTL converges to 50%; for spheres with a diameter between approximately 15 to 50 mm, RTL is typically 40%; for spheres with a small diameter (<1.2×FWHM of the PSF), RTL becomes larger than 50%.

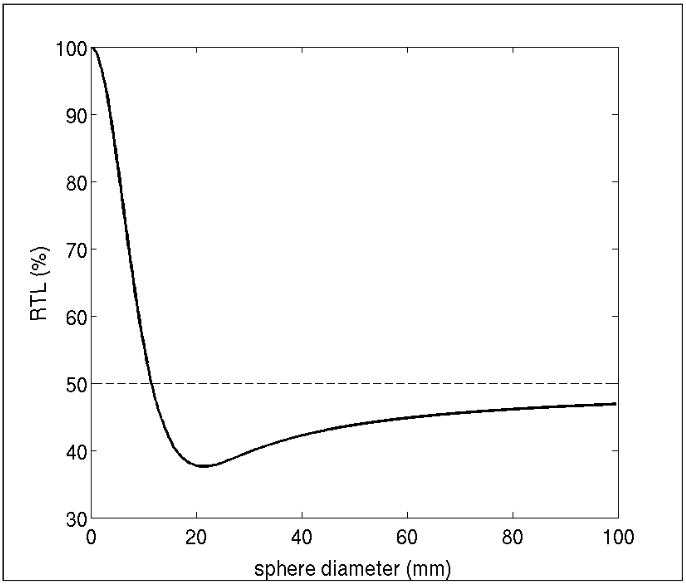


Figure 12.2
The optimal RTL as a function of the sphere diameter in case the FWHM of the point spread function (PSF) is 9.2 mm.

12.3 Materials and methods

Point spread function

The measured shape of the PSF is required to predict the optimal RTL. To obtain the PSF, an FDG point-source of 15 kBq, smaller than 2 mm in all directions, was scanned using a full-ring dedicated PET scanner (Siemens ECAT Exact 47, Siemens/CTI, USA). Emission imaging was performed for one bed position and the acquisition time was 5 minutes. The emission scan was reconstructed using 2 iterations, 8 iteration subsets and a 5 mm 3D Gaussian blurring filter. In order to achieve the most accurate measurement of the PSF, a zoom of 4 was used, resulting in a voxel size of the PET image of $1.28\times1.28\times1.28$ mm^3. The trivariate PSF was fitted with a trivariate Gaussian using the FWHM (in 3 directions) as fit parameters. Once the measured PSF is known, a convolution with a sphere can be performed, in order to obtain the optimal RTL as a function of the sphere diameter (see equation 12.4).

Phantom experiment

A phantom study was performed using a Jaszczak phantom with a volume of approximately 6.5 l, with 7 hollow spheres with inner diameters (volumes) of 10.1 mm (0.54 ml), 12.6 (1.06 ml), 15.9 mm (2.09 ml), 19.9 mm (4.15 ml), 25.0 mm (8.23 ml), 31.4 mm (16.21 ml) and 60.1 mm (113.52 ml). Activity concentration in the spheres decreased during the experiment from 25 to 6 kBq/ml; the background activity concentration was varied from 2.5 to 4.0 kBq/ml, such that SBRs were obtained in the (clinically relevant) range of 1.5 to 10.

Five FDG-PET scans of the Jaszczak phantom were acquired at 5 different SBRs: 1.8, 2.4, 3.7, 7.2 and 9.9. Similar imaging was performed as for the PSF, but including 3 minutes ^{68}Germanium/^{68}Gallium based transmission scans for attenuation correction. The voxel size of the PET images was $5.15\times5.15\times5.15$ mm^3, to match the clinical setting.

The mean background activity concentration was measured in a spherical volume of interest (VOI) of 30 mm diameter within the Jaszczak phantom, but outside the region of the spheres. In the three largest spheres, the mean activity concentration was determined by choosing a spherical VOI with a diameter 18 mm smaller (approximately 2×FHWM of the PSF) than the diameter of the sphere. It is assumed that partial volume effects (PVE) do not play a significant role for these VOIs. For the smaller spheres the activity concentration was estimated using the most intense voxel within the spheres, as opposed to a mean or median activity concentration within some VOI, which would be very sensitive to PVE. These measurements provided values for B and S.

The phantom study was performed to verify the theoretical model and to create a procedure to perform volumetric quantification of PET lesions in clinical practice. To obtain the cut-off value experimentally, we proceeded as follows. For a given sphere, voxels with values above some predefined threshold were selected. This threshold was varied until the volume of the clustered voxels correponded to the true sphere volume. This cut-off threshold (T) was translated to the RTL via $RTL = (T\text{-}B) / (S\text{-}B)$. The hypothesis that the theoretical model describes the phantom data was tested using the chi-square distribution.

As already suggested by Drever et al (18), it is important to incorporate the finite wall thickness of the spheres (1 mm plexiglass). Obviously, there is no FDG-uptake in the plexiglass. Consequently, the RTL will be reduced by this effect, depending on the SBR (18). We incorporated this wall thickness into the calculations, using a typical (average) SBR of 5.

Patient feasibility study

To test the clinical applicability of the proposed method, image series of two patients (age 61 and 66; weight 74 kg and 114 kg, respectively) with histologically confirmed colorectal cancer with liver metastases were used. Both patients underwent surgical treatment of the liver metastases, allowing intraoperative measurement of the lesion size. For both patients, CT, FDG-PET and surgery were performed within a period of one month. The maximum diameter (D_{max}) of the metastases as measured during surgery was used as the gold standard.

Prior to surgery, CT scans were acquired during inspiration breathhold using a multi-slice spiral CT scanner (Siemens Somatom VolumeZoom, Erlangen, Germany). Scanning parameters were 130 kV, 120 mAs and a slice thickness of 3 mm. In the transaxial plane, the pixel size of the CT images was 0.78×0.78 mm^2. Intravenous contrast (100 ml Xenetix-300) was applied. The portal phase contrast images were used in this study. The FDG-PET scans were acquired using the same PET scanner as mentioned before. Both patients fasted for a minimum of 6 hours before intravenous injection of 241 and 269 MBq FDG, respectively. The PET protocol consisted of a whole-body emission scan from the hips to the base of the skull, with Germanium-68/Gallium-68-based attenuation correction. The voxel size of the PET images was $5.15 \times 5.15 \times 5.15$ mm^3. The reconstruction was performed identical to the reconstruction of the phantom images.

Volumetric quantification of liver lesions on CT was performed using the software prototype OncoTREAT (MeVis, Bremen, Germany) (21-23). The underlying method is based on an automated segmentation method that is based on morphological processing. The volumes of the lesions and the maximum diameter were extracted. The gold standard was the intraoperative measurement, while CT provided a second reference for comparison with the FDG-PET results.

For FDG-PET, the volume and maximum diameter were retrieved using two different methods: a standard method (st-PET) that uses a fixed percentage of the signal S $(T/S\times 100\%)$ as a threshold level, which is similar to using a fixed SUV value and which is mostly used in clinical practice, and the RTL-based method (RTL-PET). For the standard method the volume and maximum diameter were obtained for a range of values of T. For the RTL-method, first a 40% RTL is used to get a first volume estimate. This volume V is converted to an average diameter via $D = (V\times 6/\pi)^{1/3}$. The appropriate RTL for that diameter is found using equation 12.4. This process is iterated to improve the estimate of the PET volume, until the change in the RTL value is less than 0.05%. The iteration process always converges as the RTL curve (see figure 12.2) only shows one global minimum. Spherical VOIs (with adjustable diameter) were used to select the signal S and background B: for S this volume is placed within the lesion, for B in a region of healthy liver tissue. The mean activity concentrations within these VOIs represent S and B, respectively.

12.3 Results

Figure 12.3 shows the measured PSF of the FDG point source, projected onto the transaxial plane and the result of the trivariate Gaussian fit. The FWHM fit parameters were found to be 8.9 mm (S.D. 0.1 mm), 9.3 mm (S.D. 0.1 mm) and 9.3 mm (S.D. 0.1 mm) in the x, y and z direction, respectively.

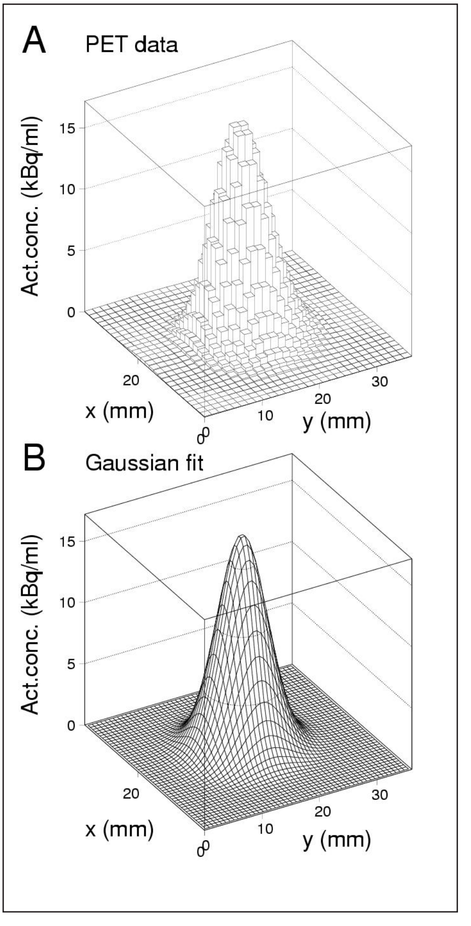


Figure 12.3
The point-spread function (PSF) of an FDG point source, projected onto the transaxial plane (A). The result of the trivariate Gaussian fit (B).

Phantom experiment

Figure 12.4 shows the optimal RTL as a function of sphere diameter as measured with the phantom. The error bars on the data points represent the standard deviation resulting from the variation observed in the 5 scans. Also shown are two calculations, based on a convolution using a trivariate Gaussian PSF with a FWHM of 9.2 mm (average result) for all 3 directions. Calculation A (which is equal to the calculation shown in figure 12.2) neglects the wall thickness of the spheres, while calculation B incorporates the wall thickness of 1 mm. As expected, calculation B best describes the data (A: $\chi^2 = 8.4$, p = 0.2; B: $\chi^2 = 3.8$, p=0.7).

The measured RTL for medium sized spheres (diameter 20-60 mm) is typically 35%, for larger spheres it slowly increases to eventually more than 40%. For smaller spheres, RTL rapidly increases with decreasing diameter. Figure 12.5 demonstrates two RTL derivations: one for a medium sized sphere (diameter 31 mm) and one for a small sphere (diameter 13 mm). Profiles in a transaxial plane and through the centre of the spheres are shown. The optimal RTL was found to be 35% and 53%, respectively. Since the value of the FWHM of the PSF approximates the diameter of the small sphere, the signal within this sphere underestimates the true activity concentration.

Patient feasibility study

An example of a liver lesion in patient 1 and segmentation as seen on both PET and CT is shown in figure 12.6. For patient 1 we obtained S = 15.5 kBq/ml (SUV=4.8) and B = 8.2 kBq/ml, resulting in SBR = 1.9. For patient 2 we obtained S = 19.2 kBq/ml (SUV=8.1) and B = 4.3 kBq/ml, resulting in SBR=4.5. Furthermore, for patient 1 we obtained RTL = 40% and for patient 2 RTL = 42%. Via $T = (S\text{-}B) \times RTL + B$, this implies an RTL-optimized threshold of T = 11.1 kBq/ml and T = 10.6 kBq/ml, respectively. This converts to a treshhold of $T/S \times 100\%$ = 72% (SUV = 3.5) and 55% (SUV = 4.5) respectively.

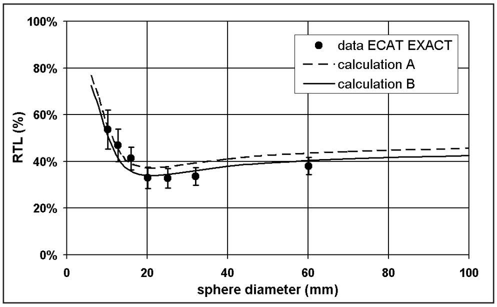


Figure 12.4
Optimal relative threshold level (RTL) as a function of the sphere diameter. The dashed curve (calculation A) is based on a convolution with a trivariate Gaussian PSF with FWHM=9.2 mm in all directions. The continuous curve (calculation B) takes into account the plexiglass wall-thickness of the spheres of 1 mm. The results of phantom measurements are well described by calculation B. For sphere diameters sufficiently large (> 1.2 × FWHM of PSF) the optimal RTL is systematically below 50%.

Figure 12.5
Illustration of two RTL derivations: one for a medium sized sphere (diameter 31 mm), A, and one for a small sphere (diameter 13 mm), B. Profiles in a transaxial plane and through the centre of the spheres are shown. The RTL was found to be 35% and 53%, respectively. Since the value of the FWHM of the point-spread function (PSF) is around the size of the small sphere, the signal within this sphere underestimates the true activity concentration.

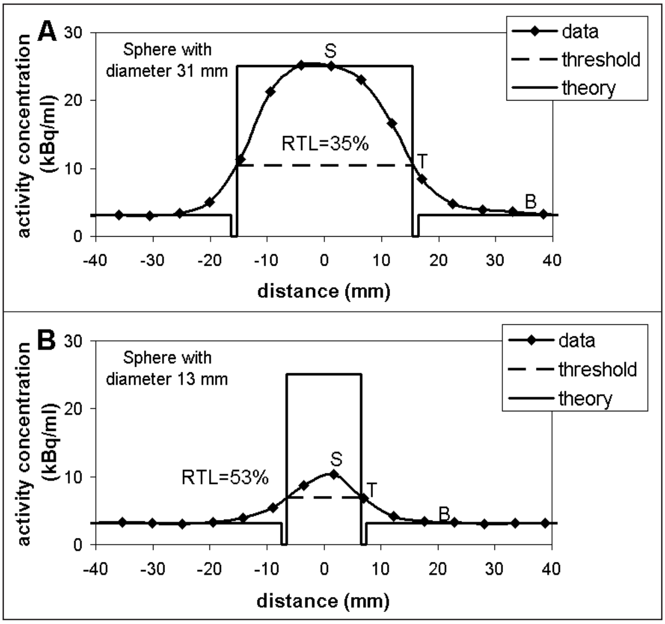


Table 12.1 summarizes all measurements of volume V and D_{max} from RTL-PET, CT and surgery, for both patients. A good agreement is observed between D_{max} from RTL-PET and surgery or CT. For st-PET, D_{max} is computed using a percentage of the signal S, where we allowed this percentage to vary; i.e. D_{max} is computed as a function of T/S. The results and those from RTL-PET, CT and surgery for both patients are shown in figure 12.7. It is clear that a fixed percentage (or a fixed SUV value) for both patients can never result in an agreement with both surgery measurements (or CT measurements). E.g., a T/S of 70%, or a threshold SUV of 3.4 (for patient 1), and a T/S of 62%, or a threshold SUV of 2.8 (for patient 2), are needed to meet the gold standard obtained from intraoperative measurement. To meet the CT results, also different values for T/S are needed: 77% for patient 1, 60% for patient 1.

	RTL-PET		**CT**		**Surgery**	
	V(ml)	D_{max} (mm)	V(ml)	D_{max} (mm)	V(ml)	D_{max} (mm)
Patient 1 (pt1)	25	39	12	31	NA	40
Patient 2 (pt2)	42	50	40	48	NA	47

Table 12.1
A summary of all volumes (V) and maximum diameters (D_{max}) of a lesion assessed by RTL-PET, CT and surgery, for both patients.

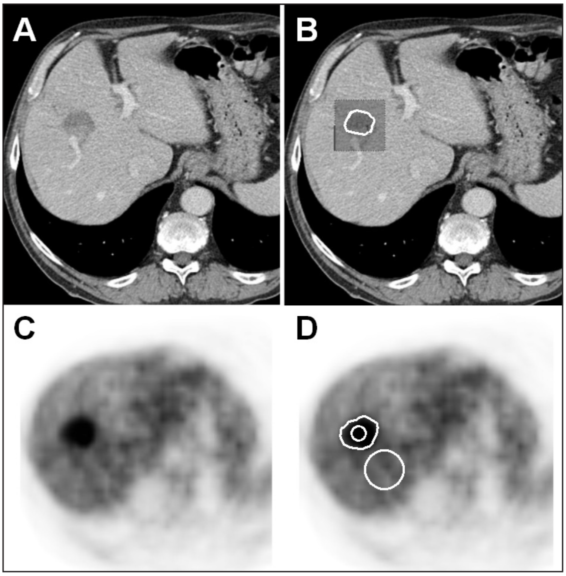


Figure 12.6
Liver lesion and segmentation as seen on a transaxial slice on both CT (A and B) and PET (C and D), for patient 1. For CT the segmentation was performed using OncoTREAT, for PET using the RTL-method. The (non-circular) contours in B and D represent the segmentation. In the PET image, signal S was determined by using a spherical VOI, represented by the small circle within the lesion in image D. The background B was determined by using a spherical VOI in a region of healthy liver tissue, represented by the large circle in image D.

Unlike the standard method, the RTL-method provides optimal threshold levels that are independent of the SBR and, therefore, this method appears to be more robust. Using RTL-PET, the difference of D_{max} with surgery or CT is small for both patients (<10%). Using st-PET, this strongly depends on the SBR. A specific T/S may easily lead to differences much larger than 10%, e.g. (with respect to surgery) 70% and 13% for patient 1 and patient 2 respectively, when a fixed threshold of 50% is used, or, 24% and 4% respectively, when a fixed threshold of 60% is used.

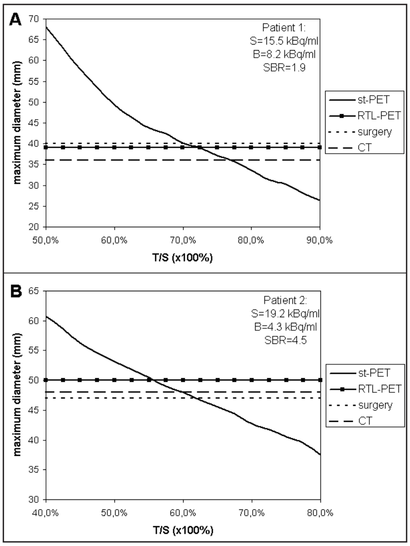


Figure 12.7
The maximum diameter as a function of T/S for both patients, using the st-PET method. Also shown are the results from RTL-PET, surgery and CT.

12.4 Discussion

Based on a convolution of a sphere with a trivariate Gaussian PSF, we derived a mathematical formalism for volumetric quantification of spheres with FDG-PET. Phantom measurements confirmed the formalism. The advantage of this formalism is that it provides optimal relative threshold levels that take into account the presence of background activity and that are independent of the SBR. The RTL-method can easily be implemented in clinical practice and be applied to other PET-radiopharmaceuticals as well. The only input that is needed is the shape of the PSF, which can be measured for any PET-scanner, for a given reconstruction method. Once the PSF is known, the RTL needed to obtain the true volume is obtained by determining the signal S (tumor uptake) and background B (healthy tissue uptake). An iteration procedure is necessary as the optimal RTL depends on a priori knowledge of the lesion diameter.

Methods that use a fixed percentage of the signal (or SUV) to obtain the contour of a PET lesion, as illustrated by others (9,11,13), imply there is no dependency on the background activity concentration B. Subsequently, this also implies it does not explicitly depend on "S-B". However, in our study we did show an explicit "S-B" dependency (for $RTL = (T\text{-}B) / (S\text{-}B)$). E.g., when assuming that the optimal RTL is 40% for a certain object, this implies an optimal cut-off T of 46% of the signal S (when case S = 10 and B = 1 kBq/ml), while a cut-off T of 70% of the signal is required (when case S = 2 and B = 1 kBq/ml or S = 10 and B = 5 kBq/ml) to obtain the correct object size. I.e. no fixed percentage of the signal (or SUV) is ever able to describe all possible situations. This was also explicitly illustrated by the patient analysis. Moreover, as our theory and phantom measurements showed that the RTL-method provides threshold levels that are independent of the SBR, this implicitly means that there is a dependency on the signal intensity. Hence, these standard methods using a fixed percentage of the signal are less suitable and, therefore, are not advisable for volumetric quantification with PET.

Other methods (14,15) do take into account the background B. However, these methods still show an SBR dependence (and neglect a diameter dependence), whereas our method was developed to provide threshold levels independent of the SBR. Under the assumption of no diameter dependence, these methods can be rewritten such that they are mathematically identical to our method. In the study of Daisne et al. (15) the relative threshold, defined as $y = T/S$, was fitted to the hyperbolic function $y = a + b/SBR$, where a and b are fit parameters. As shown by Davis et al. (17): $a = RTL$ and $b = 1\text{-}RTL$. Furthermore, the method of Black et al. (14) can also be rewritten such that it is mathematically identical to our method. Black et al. derived a linear relation between the threshold SUV (TSUV) and the signal SUV: $TSUV = c \times SUV + d$, where c and d are fit parameters, and by definition $SUV = f \times S$ and $TSUV = f \times T$, with f the body weight (in g) divided by the injected activity (in Bq). From our definition $RTL = (T\text{-}B)/(S\text{-}B)$, we can deduce $T = RTL \times S + B \times (1\text{-}RTL)$. By multiplying both sides of this equation by f we obtain $TSUV = RTL \times SUV + f \times B \times (1\text{-}RTL)$. Hence, compared with the linear relation of Black et al, we obtain the constraints $c = RTL$ and $d = f \times B \times (1\text{-}RTL)$.

The above-mentioned derivations indicate that the two previously presented methods (14,15) are based on two parameters (a and b, or, c and d, respectively), where we only used one parameter (RTL, which is equal to a and c). Hence, parameters b and d are redundant. Furthermore, these methods present a fixed value for parameter $a\ (= c = RTL)$ for a range of volumes (or diameters), whereas we showed there is a dependency. In a study by Yaremko et al. (12), the background B was taken into account, as well as a size-dependence. However, analyses were based on a fixed percentage of the signal S. Hence, an SBR-dependence was observed. Furthermore, it provided phenomenological results only, whereas we presented a model that describes the results and which can be used to estimate the appropriate threshold that will reproduce the given lesion volume most accurately. Several studies were performed where the background-subtracted relative-threshold (i.e. a fraction of "$S\text{-}B$") was used for volumetric quantification using PET. Ciernik et al (16) proposed a fixed RTL of 50%. In general this results in an underestimation and for small volumes in an overestimation of the true volume. Drever et al. (18) showed that the optimal RTL depends on the volume of the sphere. Their phantom measurements support our results. However, they did not derive the underlying theory. Hence, no prediction was (or could be) made how to implement the measurements in a clinical setting (based on a measurement of the PSF and an iteration procedure).

The FWHM of the PSF is related, but not identical, to the spatial resolution of the PET scanner. According to the NEMA NU-2 standard (24), the spatial resolution should be measured using filtered back-projection (without smoothing). However, the FWHM was measured on iterative reconstructed images, to resemble imaging in the clinical setting. Using the NEMA NU 2-2001 standard, we observed a FWHM of approximately 6 mm, significantly smaller than the FWHM of the PSF used in our analyses. I.e., our measured FWHM of the PSF is (by definition) not identical to the spatial resolution measurement according to the NEMA NU 2-2001 standard. Furthermore, the value of the FWHM of the PSF slightly depends on the position within the field of view and also in which direction it is measured. Depending on this position and the different directions, differences of typically 10% were observed within the FOV. Based on calculations, the impact on the RTL-results is expected to be small. Drever et al. (18) showed a detailed analysis of possible RTL-dependencies. Besides a dependence of the size of the object, as we also observed, they also noticed an SBR dependence, a voxel-size dependence and a slice-location dependence.

The SBR dependence is a consequence of the finite wall-thickness of the sphere. As this wall does not accumulate FDG, the presence of this wall lowers the RTL. As illustrated in figure 12.4, we also observed a wall-thickness effect. We used a typical (average) SBR of 5. However, we also performed calculations with lower and higher SBR values. In general, when the SBR increased, the wall-thickness effect decreased. However, due to the statistical fluctuations in our phantom data we could not confirm, nor reject, this tendency experimentally. In clinical practice this issue is of minor importance. In general there will be no "wall" with zero FDG-accumulation around the lesion.

In our study, the voxel-size dependence was neglected. However, an effect can be expected. E.g. if the centre of a sphere corresponds to the centre of a voxel, a different result will be obtained for the RTL than if this centre would correspond to a corner of a voxel. For sphere diameters larger than 10 mm, this effect is assumed to be negligible. This is supported by the RTL measurements that are described by our calculations over the full range of sphere sizes (10-60 mm). With an image voxel size of 5.15×5.15×5.15 mm^3, the smallest sphere-volume (D = 10 mm, V = 0.5 ml) is still about a factor 4 larger than the voxel size; for the second smallest sphere (D=13 mm, V=1.0 ml) this is already a factor of 7. Based on shifting the grid of voxels (image matrix) with respect to the centre of a sphere, such that this centre can be anywhere within a certain voxel, extra calculations confirmed that the effect of discretization into voxels is small. In the determination of the diameter using the RTL method, deviations of typically 1 mm can be expected (for D > 10 mm) with respect to the true diameter.

No slice dependency was included in our analyses. Normalization of the PET detector elements assured that fluctuations between slices were within a few percent and comparable to statistical noise. Therefore, no slice correction was needed.

When the size of the sphere approaches (or gets smaller than) the FWHM of the PSF, there is a strong increase in the optimal RTL. For small spheres, the corresponding RTL is, therefore, prone to errors (illustrated by the relatively large error bars in figure 12.4): a small change in size results in a large change in the optimal RTL. Hence, in clinical practice difficulties might be expected in volumetric quantification of small lesions. Furthermore, the RTL-method was shown to be diameter-dependent. I.e., the optimal RTL depends on the curvature of the object. Consequently, in case of non-spherical objects (e.g. ellipsoid-like objects) there is no unique RTL that can provide the contour of that object, precisely. However, for objects >15 mm the optimal RTL only slightly changes with the size (and thus with the curvature) of the object. Hence, a possible bias is expected to be small in clinical practice for lesions that are sufficiently large. Fortunately, it is the large lesions, that can extend to critical parts of the organ or to other organs, that are most interesting for accurate volumetric quantification.

As clinical PET images are acquired during (shallow) breathing, liver lesions are likely to be enlarged on PET due to the absolute displacement of the liver, which was found to be 13 mm on average (25). This will influence the choice of the RTL in order to acquire the correct lesion delineation. Obviously, the biggest influence is to be expected for small lesions as these may shift more than their own size during breathing. The possible bias being introduced is not inherent to our method. The standard lesion-delineation methods using PET will be biased in a similar way. A possible solution to deal with breathing effects is to perform respiratory gating (26).

Furthermore, in this study the anatomical location of the lesion was disregarded. It would be interesting to investigate the spatial overlap of volumes determined by the different modalities (PET and CT). Image fusion is a tool to investigate similarity measures of lesion volumes determined by PET and CT. This way, one can learn about regions identified positive on PET and negative on CT, and visa versa. Moreover, it provides an additional validation tool (27,28). Due to the difference in PET and CT breathing protocols, leading to liver deformations, we could not perform image fusion with sufficient accuracy for this specific application (29).

In clinical practice, lesions will not be spherical, nor will the accumulation be uniformly distributed within the lesion. Also the activity concentration in the background will not be uniformly distributed. Another inherent limitation is the need to manually identify the lesion and normal background and thus introduce some operator-dependency. For these reasons, the proposed method should be considered as a first approximation to perform volumetric quantification with FDG-PET, which is amendable for further improvements. However, the other published methods suffer from the same limitations. Moreover, they lack a basis on both mathematics and phantom measurements, and are not supported by a patient feasibility analysis. In particular, using a fixed threshold (e.g. 50% of the signal, or a SUV of 2.5), which is mostly done in clinical practice, may lead to errors in volumetric measurements that will not be observed with our proposed method.

12.5 Conclusion

An iterative background-subtracted relative-threshold level (RTL) method that is optimal for spherical objects of different diameters was derived, for lesion delineation and volumetric quantification with FDG-PET. This RTL-method has been shown to provide optimal threshold levels that take into account presence of background activity and are independent of the signal-to-background ratio. It is applicable in phantom and in patient studies using an iterative procedure, as the optimal RTL depends on a priori knowledge of the lesion volume (or diameter). It is a promising tool for volumetric quantification of PET lesions.

12.6 References

1. van Baardwijk A, Baumert BG, Bosmans G, van Kroonenburgh M, Stroobants S, Grégoire V, Lambin P, and de Ruysscher D. The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning. Cancer Treat. Rev. 2006;
2. Pirotte B, Goldman S, Dewitte O, Massager N, Wikler D, Lefranc F, Ben Taib NO, Rorive S, David P, Brotchi J, and Levivier M. Integrated positron emission tomography and magnetic resonance imaging-guided resection of brain tumors: a report of 103 consecutive procedures. J. Neurosurg. 2006;104:238-253
3. Buscarini E, Savoia A, Brambilla G, Menozzi F, Reduzzi L, Strobel D, Hansler J, Buscarini L, Gaiti L, and Zambelli A. Radiofrequency thermal ablation of liver tumors. Eur. Radiol. 2005;15:884-894
4. Gayowski TJ, Iwatsuki S, Madariaga JR, Selby R, Todo S, Irish W, and Starzl TE. Experience in hepatic resection for metastatic colorectal cancer: analysis of clinical and pathologic risk factors. Surgery. 1994;116:703-710
5. Blokhuis TJ, van der Schaaf MC, van den Tol MP, Comans EF, Manoliu RA, and van der Sijp, Jr.. Results of radio frequency ablation of primary and secondary liver tumors: long-term follow-up with computed tomography and positron emission tomography-18F-deoxyfluoroglucose scanning. Scand. J. Gastroenterol. Suppl. 2004;93-97
6. Podnos YD, Henry G, Ortiz JA, Ji P, Cooke J, Cao S, and Imagawa DK. Laparoscopic ultrasound with radiofrequency ablation in cirrhotic patients with hepatocellular carcinoma: technique and technical considerations. Am. Surg. 2001;67:1181-1184
7. Antoch G, Vogt FM, Veit P, Freudenberg LS, Blechschmid N, Dirsch O, Bockisch A, Forsting M, Debatin JF, and Kuehl H. Assessment of liver tissue after radiofrequency ablation: findings with different imaging procedures. J. Nucl. Med. 2005;46:520-525
8. Dromain C, de Baere T, Elias D, Kuoch V, Ducreux M, Boige V, Petrow P, Roche A, and Sigal R. Hepatic tumors treated with percutaneous radio-frequency ablation: CT and MR imaging follow-up. Radiology. 2002;223:255-262
9. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, and Kirsch CM. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J. Nucl. Med. 2005;46:1342-1348
10. Nestle U, Walter K, Schmidt S, Licht N, Nieder C, Motaref B, Hellwig D, Niewald M, Ukena D, Kirsch CM, Sybrecht GW, and Schnabel K. 18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. Int. J. Radiat. Oncol. Biol. Phys. 1999;44:593-597

11. Ciernik IF, Huser M, Burger C, Davis JB, and Szekely G. Automated functional image-guided radiation treatment planning for rectal cancer. Int. J. Radiat. Oncol. Biol. Phys. 2005;62:893-900

12. Yaremko B, Riauka T, Robinson D, Murray B, McEwan A, and Roa W. Threshold modification for tumour imaging in non-small-cell lung cancer using positron emission tomography. Nucl. Med. Commun. 2005;26:433-440

13. Wang D, Schultz CJ, Jursinic PA, Bialkowski M, Zhu XR, Brown WD, Rand SD, Michel MA, Campbell BH, Wong S, Li XA, and Wilson JF. Initial experience of FDG-PET/CT guided IMRT of head-and-neck carcinoma. Int. J. Radiat. Oncol. Biol. Phys. 2006;65:143-151

14. Black QC, Grills IS, Kestin LL, Wong CY, Wong JW, Martinez AA, and Yan D. Defining a radiotherapy target with positron emission tomography. Int. J. Radiat. Oncol. Biol. Phys. 2004;60:1272-1282

15. Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, and Grégoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother. Oncol. 2003;69:247-250

16. Ciernik IF, Dizendorf E, Baumert BG, Reiner B, Burger C, Davis JB, Lutolf UM, Steinert HC, and Von Schulthess GK. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. Int. J. Radiat. Oncol. Biol. Phys. 2003;57:853-863

17. Davis JB, Reiner B, Huser M, Burger C, Szekely G, and Ciernik IF. Assessment of (18)F PET signals for automatic target volume definition in radiotherapy treatment planning. Radiother. Oncol. 2006;80:43-50

18. Drever L, Robinson DM, McEwan A, and Roa W. A local contrast based approach to threshold segmentation for PET target volume delineation. Med. Phys. 2006;33:1583-1594

19. Kessler RM, Ellis JR, Jr., and Eden M. Analysis of emission tomographic scan data: limitations imposed by resolution and background. J Comput. Assist. Tomogr. 1984;8:514-522

20. King MA, Long DT, and Brill AB. SPECT volume quantitation: influence of spatial resolution, source size and shape, and voxel size. Med. Phys. 1991;18:1016-1024

21. Bornemann L, Kuhnigk JM, Dicken V, Zidowitz S, Kuemmerlen B, Krass S, Peitgen HO, Wein BB, Schubert H, Shin HO, and Wormanns D. Informatics in radiology (infoRAD): new tools for computer assistance in thoracic CT part 2. Therapy monitoring of pulmonary metastases. Radiographics. 2005;25:841-848

22. Kuhnigk JM, Dicken V, Zidowitz S, Bornemann L, Kuemmerlen B, Krass S, Peitgen HO, Yuval S, Jend HH, Rau WS, and Achenbach T. Informatics in radiology (infoRAD): new tools for computer assistance in thoracic CT. Part 1. Functional analysis of lungs, lung lobes, and bronchopulmonary segments. Radiographics. 2005;25:525-536

23. Kuhnigk JM, Dicken V, Bornemann L, Bakai A, Wormanns D, Krass S, and Peitgen HO. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. IEEE Trans. Med. Imaging. 2006;25:417-434

24. Ube-Witherspoon ME, Karp JS, Casey ME, DiFilippo FP, Hines H, Muehllehner G, Simcic V, Stearns CW, Adam LE, Kohlmyer S, and Sossi V. PET performance measurements using the NEMA NU 2-2001 standard. J. Nucl. Med. 2002;43:1398-1409

25. Brandner ED, Wu A, Chen H, Heron D, Kalnicki S, Komanduri K, Gerszten K, Burton S, Ahmed I, and Shou Z. Abdominal organ motion measured using 4D CT. Int. J Radiat. Oncol. Biol. Phys. 2006;65:554-560

26. Nehmeh SA, Erdi YE, Pan T, Pevsner A, Rosenzweig KE, Yorke E, Mageras GS, Schoder H, Vernon P, Squire O, Mostafavi H, Larson SM, and Humm JL. Four-dimensional (4D) PET/CT imaging of the thorax. Med. Phys. 2004;31:3179-3186

27. Daisne JF, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, and Gregoire V. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology. 2004;233:93-100

28. Riegel AC, Berson AM, Destian S, Ng T, Tena LB, Mitnick RJ, and Wong PS. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. Int. J Radiat. Oncol. Biol. Phys. 2006;65:726-732

29. Goerres GW, Kamel E, Heidelberg TN, Schwitter MR, Burger C, and Von Schulthess GK. PET-CT image co-registration in the thorax: influence of respiration. Eur. J Nucl Med. Mol. Imaging. 2002;29:351-360

13

Delineation and measurement of liver metastases on FDG-PET, with pathological or intra-operative verification

Wouter V. Vogel [1]
Jorn A. van Dalen [1]
Bas Wiering [2]
Heleen Dekker [3]
Volker Dicken [4]
Frans H.M. Corstens [1]
Theo J.M. Ruers [2]
Wim J.G. Oyen [1]

Departments of Nuclear medicine [1], Surgery [2], and Radiology [3],
Radboud University Nijmegen Medical Centre, the Netherlands.
MeVis-center for medical diagnostic systems and visualisation [4], Bremen, Germany.

Submitted.

Abstract

Locoregional therapies for liver metastases, such as partial liver resection (PLR) and radiofrequency ablation (RFA), require accurate delineation of malignant tissue. This is traditionally provided by CT imaging. Functional imaging with FDG-PET provides an additional tool for visualization of liver metastases, but suffers from uncertainties due to the relatively poor spatial resolution and not-straightforward lesion delineation. The aim of the study was to determine an optimized delineation procedure for size measurement of liver metastases on PET.

Methods: Studied were thresholds for PET based on multiple standard uptake values (SUV), different percentages of the maximum lesion intensity (LI) and average background intensity (BI), and an iterative relative thresholding level (RTL) technique which corrects for the influences of lesion size and background intensity. First, lesions were delineated in a phantom with hot spheres with known sizes, at different lesion-to-background ratios. Second, liver metastases were delineated on FDG-PET and CT in patients with colorectal carcinoma, prior to laparotomy. The true largest lesion diameters were determined at pathological examination or with combined intraoperative ultrasound and palpation, for PLR specimens and RFA interventions respectively.

Results: Phantom lesions could best be measured on PET using the RTL method, with an average error of 5 mm (range 2 – 10 mm). In the patient study, 7 metastases were resected and 6 were treated with RFA. True largest diameters ranged from 15 to 80 mm. Several SUV, BI and BL thresholds failed to generate adequate delineation in one or more cases, due to physiological background activity in the liver. Successful delineation could be achieved in all cases using SUV=4.0, LI=60%, BI=140%, RTL, and CT, with average errors in the measured largest lesion sizes of 9, 7, 8, 7, and 6 mm, respectively. There was no significant difference between any of these successful methods.

Conclusions: The size of liver metastases from colorectal carcinoma can be determined using PET, with an accuracy in the range of the image resolution, and comparable to CT results. Multiple lesion delineation methods for PET are suitable, of which the operator-independent RTL method is preferred. The well-known thresholds SUV=2.5 and LI=50% are inadequate for this specific application.

13.1 Introduction

Accurate imaging and delineation of liver metastases is gaining importance as more therapeutic options are becoming available, such as partial liver resections and radiofrequency ablation (RFA). These locoregional interventions rely on accurate information about localization in relation to important structures such as blood vessels, and the exact size and shape of tumor sites (1,2).

Multiple imaging modalities are available for detection of tumor tissue in the liver. Anatomical imaging with computed tomography (CT) provides tumor localization and anatomical reference with high spatial resolution, but the technique may suffer from a suboptimal sensitivity and specificity for liver lesions despite the application of intravenous contrast media, especially after previous therapeutic interventions (3,4). Furthermore, uncertainties in the determination of lesion borders and size measurements may exist, because changes in tissue density and perfusion are also influenced by benign phenomena such as peritumoral edema.
Functional imaging with ^{18}F-fluor-deoxy-glucose positron emission tomography (FDG-PET) can detect tumor tissue by visualization of glucose metabolism, with a high contrast resolution between benign and malignant tissues (5). FDG-PET may be of value for tumor detection, delineation and measurement, especially where anatomical imaging is cumbersome, i.e. in soft tissues or in altered anatomy after surgical procedures (6-8). The benefit of joined capabilities of CT (anatomical reference) and FDG-PET (sensitive tumor detection) have led to the clinical practice of correlation of images as obtained by PET and by CT, with synergistic results (8-10).

Delineation and measurement of lesions is not straightforward in PET imaging. As the spatial resolution is relatively low, images tend to be somewhat blurry or noisy, and anatomical orientation is limited (11). The determination of the “true” contour of FDG accumulation is addressed in different ways by various groups. Examples of delineation on PET in radiation oncology planning include visual interpretation (12,13), contouring based on various thresholds such as a percentage of the maximum lesion intensity (14) or calculated SUV levels (15), or contouring based on ratios between tumor and background intensities (16,17). No consensus threshold exists for any of these methods, because they are all influenced by multiple factors, such as tumor size and shape, tumor metabolism, and surrounding background tissue metabolism. Specific experience in delineation of liver lesions, where physiological background activity is high, is limited. We have recently introduced a mathematically based method for volumetric quantification of PET lesions, independent of the tumor-to-background ratio, and adapted to blurring effects related to lesion size and scanner characteristics. This relative threshold level (RTL) procedure has been described in detail elsewhere (18). The choice of

method and threshold level may have significant impact on measurement and delineation of lesions, as illustrated in figure 13.1. It is currently unknown which method performs best in delineation of liver metastases, and how this performance relates to CT-based contouring.
Here, the potential of different PET delineation methods for measurement of liver metastases is evaluated, in comparison with CT, with pathology or intra-operative verification as a gold standard.

13.2 Methods and materials

Phantom experiment
The phantom experiment was conducted to determine the impact of lesion intensity, lesion size, and background intensity on delineation and maximum diameter measurements of lesions, in clinically relevant ranges. A 6.5 liter Jaszczak phantom, with 7 hollow spheres with inner diameters of 10.1, 12.6, 15.9, 19.9, 25.0, 31.4, and 60.1 mm, was used. The activity concentration in the spheres decreased during the experiment from 25 to 6 kBq/ml, during which the background activity concentration was varied from 2.5 to 4.0 kBq/ml. Thus, 5 subsequent scans could be acquired with signal to background ratios (SBR) 8.4, 7.2, 3.7, 2.4, and 1.8. The thickness of the plexiglass wall around the hot spheres was 1 mm. Hybrid PET/CT imaging parameters were identical to the clinical experiment described below.

Patients			**Lesions**					**Imaging intervals**	
	Age (Y)	M/F	Lesion	Liver segment	Lesion shape	Treatment	Real max. diam. (mm)	PET-CT (days)	PET-oper. (days)
A	60	F	1	III	Irregular	Resection	70	6	7
B	67	M	2	VI	Irregular	Resection	36	7	28
			3	IVA	Sphere	RFA	25		
C	48	M	4	IVA	Sphere	Resection	32	15	29
D	56	M	5	VI	Sphere	Resection	30	19	2
E	66	M	6	IVA	Irregular	Resection	70	5	22
F *	63	M	7	II,III,IV	Irregular	Resection	80	9	14
G *	66	M	8	IVB	Irregular	RFA	40	0	35
H *	77	M	9	I	Irregular	RFA	15	13	15
I *	66	F	10	III	Sphere	RFA	27	7	21
			11	VII	Sphere	RFA	17		
			12	VIII	Sphere	RFA	37		
J *	48	F	13	V	Sphere	Resection	25	7	28

Table 13.1
*Patients and lesion characteristics. * = PET images acquired on integrated PET/CT. RFA = Radio frequency ablation. The real maximum diameter of lesions was determined at pathological examination (for resected lesions) or intra-operative ultrasound and palpation (RFA lesions).*

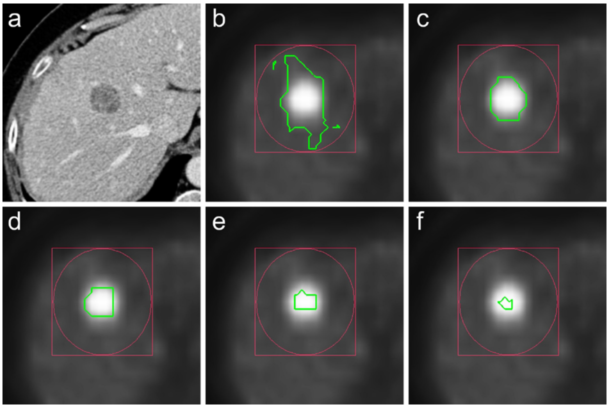


Figure 13.1
Dependancy of measured maximum lesion diameter on the chosen threshold for delineation. Lesion number 12 on CT (a) and FDG-PET with contours (green) based on SUV threshold 3.0 (b), 4.0 (c), 5.0 (d), 7.5 (e), and 10.0 (f), respectively. In general, a higher threshold results in a smaller measured lesion size. The graph also illustrates that multiple thresholds may generate visually comprehensive delineation (images c, d, e), while actually deviating significantly from the true lesion size. Other thresholds do not render adequate delineation (images b, f). In this specific case, SUV threshold 4.0 (c) correlated best with pathology findings.

The (mean) background intensity (BI) concentration at each ratio was measured in a spherical volume of interest of 30 mm diameter within the Jaszczak phantom outside the region of the spheres. The lesion intensity (LI) was defined as the activity in the hottest voxel. Threshold-based delineation was performed for all spheres at steady SBR (3.7), and for all SBR's at steady sphere diameter (31.4 mm), being representative for the clinical series that was performed. Delineation was performed using 60% of the maximum lesion intensity (LI_{60}), 140% of the average background intensity (BI_{140}), and the RTL (relative threshold level) method, as described below in the clinical experiment. These thresholds were the optimal methods for clinical evaluation, as will become apparent later on. Every voxel within the VOI with an intensity above the applied threshold was regarded pathological. From these voxels a 3D volume was rendered, from which the maximum diameter was mathematically derived.

Clinical experiment

Ten consecutive patients (3 female, 7 male; mean age 62 years, range 48 – 77 years) with known liver metastases from colorectal carcinoma that were visible on both FDG-PET and CT, and who were considered candidates for curative locoregional intervention, were studied. Thirteen liver lesions were available for analysis, of which 7 were resected and 6 were treated with RFA. Patient and tumor characteristics are summarized in table 13.1.

Imaging parameters

Prior to laparotomy all patients were imaged with dedicated contrast-enhanced CT and with FDG-PET, either as a stand-alone modality or as an integrated PET/CT. A maximum interval of 5 weeks (35 days) was accepted between imaging and surgery (see table 13.1).

Multi-detector row CT scan of the liver was performed: 4 patients using a Siemens Somatom Volume Zoom, 5 patients using a Siemens Sensation 16 and 1 patient using a Siemens Sensation 64 (Siemens, Erlangen, Germany). The scan parameters were 120 kV and 130-200 mAs. The beam collimation used was respectively 4 x 2.5 mm, 16 x 1.5 mm and 64 x 1.2 mm. Reconstructed section thickness was 3 mm. A non-contrast liver scan was followed by three distinct enhancement phases (arterial, portal and late venous phase) after administration of 150 ml Xenetix (iodine 300 mg/ml; Guerbet). Portal venous phase images were selected for evaluation and lesion delineation.

Dedicated FDG-PET scans were acquired using a Siemens ECAT Exact 47 PET-scanner (Siemens/CTI, Knoxville, Tennessee, USA). A 3D emission scan of the upper abdomen was acquired during free breathing, 60 minutes after intravenous injection of 250 MBq FDG. A 2D Germanium-68 based transmission scan was acquired for attenuation correction. The acquisition time per bed position was 5 minutes for emission and 3 minutes for transmission. Scans were reconstructed using the iterative OSEM algorithm with 2 iterations and 8 subsets, and with a 6 mm 3D Gaussian filter.

The most recent FDG-PET scans were acquired using a Siemens Biograph Duo hybrid PET/CT scanner (Siemens/CTI, Knoxville, Tennessee, USA), see table 13.1. For these scans the emission acquisition time was 4 minutes per bed position. Attenuation correction was based on low-dose CT images acquired during unforced expiration breathhold. Image acquisition and reconstruction were otherwise identical to dedicated PET imaging.

Image analysis

On PET images, liver lesions were delineated using in-house developed software for image viewing and analysis, based on the visualization toolkit VTK (19) and the insight segmentation and registration toolkit ITK (20). For analysis, each solitary liver lesion was isolated in an ellipsoid volume of interest (VOI), containing the lesion and an abundant margin of surrounding normal tissue. Lesion contours were determined within the VOI using multiple thresholding algorithms, based on different relations between lesion signal activity, lesion size, and background signal activity:

- Standard uptake value (SUV): The SUV value was calculated for all voxels, using the formula *MeasuredActivityConcentration * Bodyweight / InjectedActivity*, with correction for decay to the time of injection. For thresholding, SUV levels 2.0, 2.5, 3.0, 4.0, and 5.0 (denoted as $SUV_{2,0}$ – $SUV_{5,0}$) were chosen.

- Lesion intensity (LI): The intensity of the most active voxel in the lesion was defined as 100%. Multiple levels below 100% were chosen for separate thresholding, ranging from 20% – 90%, at 10% intervals (denoted as LI_{20} – LI_{90}).
- Background intensity (BI): The average background intensity (measured in a manually defined 3-dimensional ellipsoid VOI in visually normal liver tissue, with a diameter of ~50 mm) was defined as 100% (BI_{100}). Multiple levels above BI_{100} at 20% intervals were used for thresholding (denoted as BI_{120} – BI_{200}).
- Relative threshold level (RTL): As described earlier (18), the lesion contours were iteratively determined using a background-subtracted threshold adapted to lesion size. The relation between lesion size and optimal threshold, valid for both PET systems being used, is shown in figure 13.2. First, the average background intensity (liver) and maximum lesion intensity were determined. Second, the lesion volume was estimated using an initial threshold level of 40% between background and lesion maximum. Based on the estimated volume, the threshold level was adapted to a theoretically better suitable value, as predicted by the associated border curve. Using the new threshold, the lesion volume was again estimated. This step was iterated 4 times, to converge to a final threshold.

Identical to the phantom experiment, a 3-dimensional volume was rendered subsequently from the delineation result, from which the maximum lesion diameter was mathematically derived. A delineation method was considered "successful" when >90% of the scanned lesions could be contoured with a result that was not clearly abject upon visual inspection.

On CT images, liver lesions were delineated using OncoTREAT (MeVis, Bremen, Germany) (21). The software allows automatic contour detection of lesions within the liver based on differences in tissue density and contrast enhancement. Identical to analysis of PET images, detected lesion contours were used to generate a 3-dimensional surface map, from which the maximum lesion diameter and lesion volume were mathematically derived.

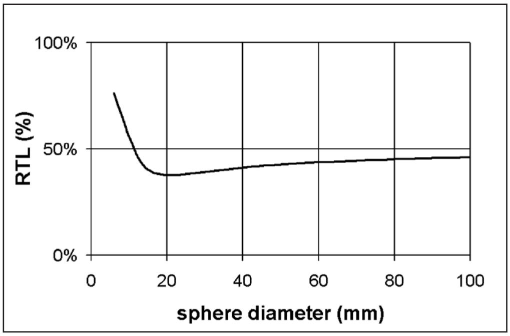


Figure 13.2
The relative threshold level (RTL) denotes the theoretically optimal threshold level for a lesion of a given size. The shape of the curve is explained by the convolution of the sphere and the measured point spread function. Because the size of a lesion is unknown, the optimal position at the curve is identified iteratively for each separate lesion.

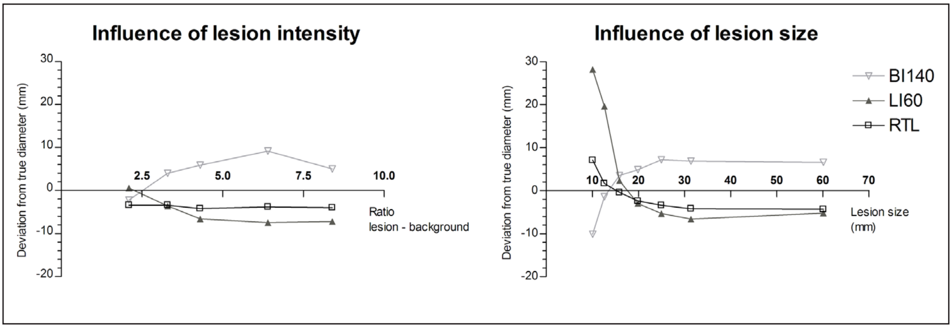


Figure 13.3
The influence of lesion intensity (for lesion size=31.4 mm) and lesion size (for SBR=3.7) on several methods for measurement of lesion diameters. The BI_{140} *and* LI_{60} *methods may suffer from extreme deviations, especially at low SBRs and for small lesions. The RTL method eliminates the influence of the intensity ratio between lesion and background, and reduces the impact of the lesion size on measurement results. A small deviation remains even for RTL measurements, due to the artificial 'cold' border between lesion and background that is present in the phantom model.*

Verification of lesion size

For patients undergoing partial liver resection, the diameter of lesions was evaluated by pathological examination. Resected specimens were subjected to standard slicing in 3 mm sections, parallel to the largest diameter as determined by palpation. The maximum diameter was measured in the plane with the largest tumor section. For all other patients, the largest diameter of lesions was determined by consensus interpretation of intra-operative ultrasound by an experienced radiologist, and palpation by an experienced liver surgeon. Unfortunately, neither pathological examination nor intra-operative ultrasound could reliably establish the total tumor volume.

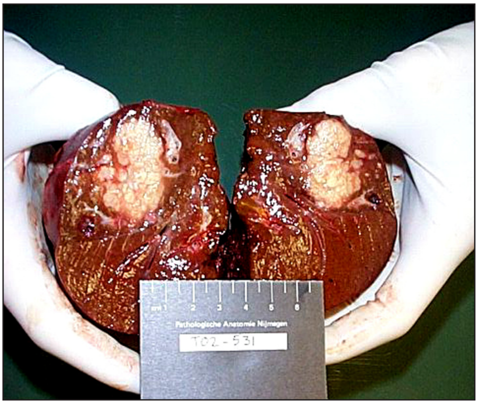


Figure 13.4
Longitudinal slice through a liver metastasis for measurement of the largest diameter. Measurements are complicated by the irregular shape of the lesion. Furthermore, slight deformation of these soft tissues is likely to occur because of handling pressure, as illustrated by the apparent difference in size of the both lesion halves.

Statistical analysis

Differences in the largest diameter of liver lesions between various delineation methods on PET, contrast-enhanced CT, and pathological or intra-operative verification, were calculated using Bartlett's test, with the level of significance set at 0.05.

13.3 Results

Phantom experiment

The results of the phantom experiment are depicted in figure 13.3. All phantom measurements suffered from a systematical bias, that was attributed to the influence of the cold 1 mm wall of the spheres (22). As this phenomenon does not play a role in clinical imaging, it was not further evaluated. The true inner diameter of the 31.4 mm sphere was reliably reproduced at all SBRs by the RTL method, when disregarding the systematical bias (an underestimation of ~4 mm on average). The LI_{60} and BI_{140} were both less accurate and showed strong dependency of the SBR, with deviations of the real sphere diameter up to –7.2 mm and +9.2 mm respectively, including the systematical bias. The spectrum of sphere diameters could not be reproduced reliably by either the LI_{60} or the BI_{140} methods. Extreme deviations occurred especially in the range of the smaller spheres, with errors up to –10.1 and +28.1 mm for LI_{60} and BI_{140} respectively. The RTL method was also influenced by the sphere size, due to the increasing relative impact of the wall thickness on the bias for smaller spheres, but clearly showed less dependency.

Clinical experiment

Thirteen liver lesions were evaluated, with an average maximum diameter of 39 mm (range 15 – 80 mm) at pathological or intra-operative verification. An example of a liver metastasis measured at pathological evaluation is shown in figure 13.4. The average time interval between PET and CT imaging was 7 days (range 0 – 35 days), and between PET imaging and surgery 20 days (range 2 – 35 days).

A clear relation was found between the chosen threshold and the measured tumor size, for all SBR-dependent methods (i.e. based on SUV, LI and BI). Lower thresholds carried a risk for inclusion of large volumes of normal liver tissue, while higher thresholds frequently excluded a significant part of the tumor volume. This relation varied among lesions and was unpredictable (figure 13.5).

SUV based evaluation frequently failed. Only $SUV_{4,0}$ could produce a successful result in >90% of cases. Besides $SUV_{4,0}$, the methods LI_{50}, LI_{60}, LI_{70}, BI_{120}, BI_{140}, BI_{160}, and RTL were also successful; further statistical analysis was performed on these methods (see table 13.2). The best results were achieved using methods LI_{60}, BI_{140}, and RTL, all with 100% success rate, an

			Measured maximum lesion diameter, deviation from gold standard		
Protocol	Succes rate (%)		Absolute average (mm)	Relative range (mm)	S.D. (mm)
CT	100	*	7	-8 – 19	6
PET SUV$_{2,0}$	0				
PET SUV$_{2,5}$	31				
PET SUV$_{3,0}$	62				
PET SUV$_{4,0}$	100	*	9	-30 – 16	13
PET SUV$_{5,0}$	77				
PET LI$_{20}$	15				
PET LI$_{30}$	46				
PET LI$_{40}$	77				
PET LI$_{50}$	92	*	7	-12 – 16	8
PET LI$_{60}$	100	*	8	-15 – 10	8
PET LI$_{70}$	100	*	10	-30 – 7	9
PET LI$_{80}$	54				
PET LI$_{90}$	15				
PET BI$_{120}$	92	*	12	-4 – 21	8
PET BI$_{140}$	100	*	7	-10 – 19	8
PET BI$_{160}$	100	*	8	-21 – 16	11
PET BI$_{180}$	85				
PET BI$_{200}$	62				
PET RTL$_{iterative}$	100	*	7	-21 – 9	8

Table 13.2 *Delineation results. * = Techniques that succeeded in delineating more than 90% of lesions with a visually acceptable result, and were thus suitable for further statistical analysis.*

absolute average deviation from the gold standard <10 mm, and a standard deviation between measurements <10 mm. None of these methods proved significantly better than any other (Bartlett’s test P > 0,1 in all combinations).

The RTL based method showed a tendency to perform better with spherical lesions, as compared with irregularly shaped lesions (average deviation range –12 – +1 mm, S.D. 5 mm, versus range –21 – +9 mm, S.D. 11 mm, respectively), although this difference was not significant (Bartlett’s test P=0,08). A similar influence of spherical lesions was seen for all delineation methods, as illustrated in figure 13.6.

CT based evaluation was successful in 100% of cases, and showed an absolute average deviation from the gold standard of 7 mm, and a standard deviation between measurements of 6 mm.

13.4 Discussion

The results of the current study validate the use of maximum tumor diameter determination on PET images using the RTL method in clinical practice, theoretically independent of the tumor-to-background ratio, correcting for the lesion size, and independent of observer factors (besides definition of a normal tissue background region), with a success rate of 100%. We consider the accuracy of the measurements adequate, as the average error of 7 mm approximates the image resolution that can be achieved in the abdominal region using current PET scanners, and is comparable to the results achieved with state-of-the-art CT-based delineation. It is likely that volumetric quantification using RTL is also adequate, but this could not be confirmed in the current study design due to lack of a gold standard for volumetric analysis.

The other methods for tumor delineation on PET (i.e. SUV, LI, and BI) were all influenced by the threshold that was chosen, thus rendering these methods observer-dependent. The fixed threshold $SUV_{2,5}$, that is often used to attempt discrimination of malignant and benign tissues $^{(23)}$, proved unsuitable for size measurement of liver lesions in this study, with a success rate of only 31%. This is caused by the relatively high background activity in normal liver tissue, which was in the range of SUV ~2 in most patients. For SUV-based delineation, a threshold higher than 2.5 for discrimination of benign and malignant lesions in the liver has been advised before, e.g. 3.5 by Delbeke et al. $^{(24)}$. In our setup, $SUV_{4,0}$ yielded reasonable results, but with a larger spread than other methods. Furthermore, calculated SUVs may vary considerably between different scanners, or even between repeated measurements of the same patient $^{(25-27)}$. We consider SUV-based methods suboptimal for evaluation of liver metastases.

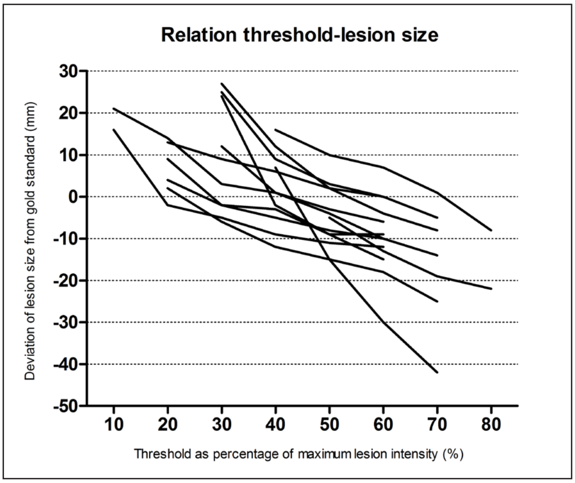


Figure 13.5
Relation between the measured maximum lesion diameter and the chosen threshold for delineation. A higher threshold results in a smaller measured lesion size. However, this relation is unpredictable and lesion-specific, depending on factors as lesion intensity, size and shape, and background intensity. No single threshold can accurately delineate all lesions.

The methods LI_{60} and BI_{140} both achieved the same accuracy as RTL in this series. Where the RTL technique is not available, these methods may be used as an adequate alternative. Nevertheless, on theoretical grounds, the RTL technique seems preferable because a better result is predicted for relatively large and relatively small lesions (18). The other BI and LI based thresholds were suboptimal, due to a low success rate or a larger spread in measurement results.

It is of concern that many of the rejected methods could produce visually reasonable lesion contours, while deviating significantly from the true lesion size. For many lesions, visually comprehensible delineation could be provided using rather extreme thresholds such as LI_{30} and LI_{80} (non-zero success rates in table 13.2). Thus, mere visual assessment is a suboptimal instrument for recognition of valid lesion delineation. Furthermore, widely-used methods could fail entirely in generating a comprehensible delineation for a specific lesion. For example, the widely used 50% contouring (LI_{50}) method failed in one case, and the $SUV_{2,5}$ method failed in 9 cases out of 13. These issues once again emphasize the need for methods that do not depend on observer, lesion, or background factors.

Effect of lesion shape

The RTL method was designed for accurate evaluation of spherical lesions, based on correction of point-spread blurring effects at the curved border of a sphere. This explains the good results for spherical lesions, as compared to irregular shaped lesions (average absolute error 5 mm versus 13 mm). In our experience with PET, the occurrence of spherical liver metastases is approximately equal to those with irregular growth and central necrosis (7 out of 13 spherical in this study). A further improvement for evaluation of irregular lesions can be expected when the RTL method can be adapted to correct for local differences in border curvature.

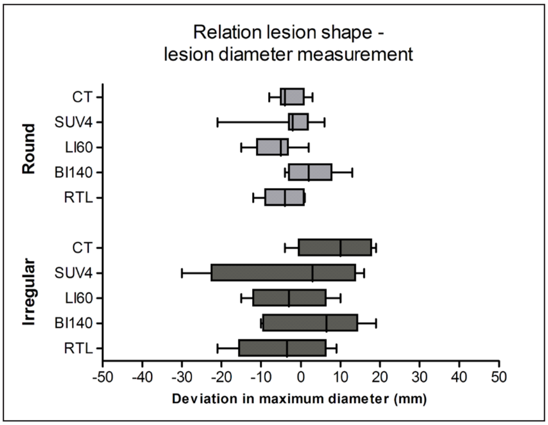


Figure 13.6
Regardless of the applied measurement method, size measurement appeared more accurate for sphere-like lesions than for irregularly shaped lesions. This is illustrated by larger deviations from the gold standard in maximum diameter measurements, for irregularly shaped lesions.

Variances in metabolic rate within a single lesion pose an additional problem. Currently no method exists to discriminate inhomogeneous lesions from irregularly shaped lesions. This will continue to compromise lesion delineation using PET. The RTL method is therefore recommended for spherical lesions and homogeneous lesions, while lesion size definition for inhomogeneous lesions has to be cautiously interpreted using any method.

Micrometastases (<1 cm) form a separate category. In general, FDG-PET is not adequate for detection of micrometastases in the liver, due to the relatively poor image resolution and relatively high background uptake in normal liver tissue (11,28). When small lesions are detected, signal intensity may be only marginally higher than the background, thus limiting the possibilities for thresholding. On theoretical grounds, delineation and volumetric quantification of – sufficiently intense – small lesions may theoretically best be performed using the RTL method (18). In small lesions with very low intensity, and lesions approaching the voxel size, attempts to delineate on PET are discouraged.

Imaging

In this study, PET delineation and size measurements were compared with CT imaging. It is known that CT is not flawless with regard to detection of liver metastases, despite the application of intravenous contrast (29,30). However, CT imaging does represent the current clinical standard of care. The fact that CT and PET performed similar at size measurement of liver metastases, when an adequate PET delineation method was used, supports the conclusion that clinical value can be derived from both imaging modalities. Uncertainties and inaccuracies may have been introduced in this study due to a change of equipment. Early PET scans were acquired using a dedicated PET camera, later scans using an integrated PET/CT scanner (table 13.1). Although some imaging characteristics varied among the scanners (e.g. voxel size, detector efficiency), the point-spread-function and effective image resolution (FWHM) were similar. When using an entirely different scanner, or e.g. a different image reconstruction protocol, the RTL method remains valid as long as the point-spread function is adjusted. Similar effects may also apply for the LI and BI-based methods. It may be advised to determine the characteristics of a specific camera, and validate size measurements, before using any kind of lesion delineation. In this study, no significant differences could be demonstrated between the results achieved with the dedicated PET and integrated PET/CT scanners (data not shown).

Limitations of the approach

Determination of the maximum lesion size in pathological examination is not trivial. The choice of the slicing plane was based on palpation (ex vivo) and may have been suboptimal. Small tumor bulges may have existed just off the slicing plane. Nevertheless, this technique was considered the approach closest to a gold standard. Complete volumetric analysis of excised lesions was considered technically not feasible.

The alternative of intra-operative ultrasound combined with intra-operative palpation, in consensus by an experienced radiologist and an experienced surgeon respectively, may be slightly less accurate as compared to pathological examination. Some lesions were located in the central (deep) part of the liver, compromising both palpation and ultrasound, and further bias may be caused by extensive peri-tumoral edema (31). Some have applied CT imaging as the gold standard (32), although we feel that evidence for such an approach is limited. Our results do demonstrate good correlation between PET and CT measurements, as well as between CT and pathology measurements.

Time interval

Liver metastases may be rapidly progressive in size and number, or may develop central necrosis over time. This implies that a short time interval between different imaging modalities and operative verification is required for adequate correlation of results. In this study, the intervals between imaging modalities (average 9 days, max. 35 days), and between FDG-PET and surgery (average 20 days, max. 35 days) were considered reasonable. Nevertheless, tumor growth in the interval between imaging and surgery may in part account for the systematical underestimation of lesion sizes (4 mm on average when using the RTL measurements on PET). A prospective study could overcome this issue, if a repeated PET and CT can be incorporated directly prior to surgery.

It can be concluded that the size of liver lesions can be determined accurately using PET and the RTL algorithm, with an average measurement error in the range of the image resolution of the scanner, which is comparable with CT-based delineation. It is uncertain whether a higher accuracy can be determined using current PET scanners. No diagnostic technique can provide size measurements or delineation as an absolute gold standard, because the presence of a tumor capsula, microscopic tumor bulges and satelites, peri-tumoral edema, reactive inflammation, and changes in vascularization, all influence evaluations of any kind. The current accuracy achieved with PET may suffice for clinical implementation, i.e. planning of partial liver resections or RFA procedures. A further benefit would be needed not from improved size measurement, but from better localization and orientation in relation to anatomical structures (e.g. blood vessels and intrahepatic bile ducts).

For localization purposes, multimodality imaging may be of added value. Vascular structures can be identified on anatomical imaging such as CT and MRI, but not on FDG-PET. On the other hand, PET imaging is appreciated for its sensitive tumor detection, even in an anatomically deformed liver after previous therapeutic interventions (6,7). The added value of integrated PET/CT imaging for liver lesions has been widely recognised (8,33,34), although breathing motion differences in the region of the diaphragm may be of relevance (35). Fused image sets of liver lesions are not yet perfect in the current implementations of PET/CT, as well as in other combinations such as PET/MRI. Currently, estimation of the size of liver lesions using RTL based analysis of PET imaging in combination with co-localization relative to vessels as provided by CT or MRI seems the best that can be achieved. More research and development of improved image registration techniques seems a prerequisite for further clinical implementation of multimodality evaluation of liver metastases.

13.5 Conclusions

The size of liver metastases from colorectal carcinoma can be determined accurately using PET imaging and the RTL algorithm, with an average measurement error in the range of the image resolution of the scanner, and comparable with measurements on CT images. When RTL is not available, lesion contouring using 60% of the maximum lesions intensity or 140% of the background intensity seem adequate alternatives.

13.6 References

1. Buscarini E, Savoia A, Brambilla G, Menozzi F, Reduzzi L, Strobel D, Hansler J, Buscarini L, Gaiti L, and Zambelli A. Radiofrequency thermal ablation of liver tumors. Eur. Radiol. 2005;15:884-894
2. Gayowski TJ, Iwatsuki S, Madariaga JR, Selby R, Todo S, Irish W, and Starzl TE. Experience in hepatic resection for metastatic colorectal cancer: analysis of clinical and pathologic risk factors. Surgery. 1994;116:703-710
3. Dromain C, de BT, Elias D, Kuoch V, Ducreux M, Boige V, Petrow P, Roche A, and Sigal R. Hepatic tumors treated with percutaneous radio-frequency ablation: CT and MR imaging follow-up. Radiology. 2002;223:255-262
4. Lim HK, Choi D, Lee WJ, Kim SH, Lee SJ, Jang HJ, Lee JH, Lim JH, and Choo IW. Hepatocellular carcinoma treated with percutaneous radio-frequency ablation: evaluation with follow-up multiphase helical CT. Radiology. 2001;221:447-454
5. Weber WA, Schwaiger M, and Avril N. Quantitative assessment of tumor metabolism using FDG-PET imaging. Nucl. Med. Biol. 2000;27:683-687
6. Antoch G, Vogt FM, Veit P, Freudenberg LS, Blechschmid N, Dirsch O, Bockisch A, Forsting M, Debatin JF, and Kuehl H. Assessment of liver tissue after radiofrequency ablation: findings with different imaging procedures. J. Nucl. Med. 2005;46:520-525
7. Langenhoff BS, Oyen WJ, Jager GJ, Strijk SP, Wobbes T, Corstens FH, and Ruers TJ. Efficacy of fluorine-18-deoxyglucose positron emission tomography in detecting tumor recurrence after local ablative therapy for liver metastases: a prospective study. J. Clin. Oncol. 2002;20:4453-4458

8. Selzner M, Hany TF, Wildbrett P, McCormack L, Kadry Z, and Clavien PA. Does the novel PET/CT imaging modality impact on the treatment of patients with metastatic colorectal cancer of the liver? Ann. Surg. 2004;240:1027-1034

9. Blokhuis TJ, van der Schaaf MC, van den Tol MP, Comans EF, Manoliu RA, and van dS, Jr.. Results of radio frequency ablation of primary and secondary liver tumors: long-term follow-up with computed tomography and positron emission tomography-18F-deoxyfluoroglucose scanning. Scand. J. Gastroenterol. Suppl. 2004;93-97

10. Veit P, Antoch G, Stergar H, Bockisch A, Forsting M, and Kuehl H. Detection of residual tumor after radiofrequency ablation of liver metastasis with dual-modality PET/CT: initial results. Eur. Radiol. 2005;

11. Ruers TJ, Langenhoff BS, Neeleman N, Jager GJ, Strijk S, Wobbes T, Corstens FH, and Oyen WJ. Value of positron emission tomography with [F-18]fluorodeoxyglucose in patients with colorectal liver metastases: a prospective study. J. Clin. Oncol. 2002;20:388-395

12. Nestle U, Walter K, Schmidt S, Licht N, Nieder C, Motaref B, Hellwig D, Niewald M, Ukena D, Kirsch CM, Sybrecht GW, and Schnabel K. 18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. Int. J. Radiat. Oncol. Biol. Phys. 1999;44:593-597

13. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, and Kirsch CM. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-Small cell lung cancer. J. Nucl. Med. 2005;46:1342-1348

14. Ciernik IF, Huser M, Burger C, Davis JB, and Szekely G. Automated functional image-guided radiation treatment planning for rectal cancer. Int. J. Radiat. Oncol. Biol. Phys. 2005;62:893-900

15. Wang D, Schultz CJ, Jursinic PA, Bialkowski M, Zhu XR, Brown WD, Rand SD, Michel MA, Campbell BH, Wong S, Li XA, and Wilson JF. Initial experience of FDG-PET/CT guided IMRT of head-and-neck carcinoma. Int. J. Radiat. Oncol. Biol. Phys. 2006;65:143-151

16. Black QC, Grills IS, Kestin LL, Wong CY, Wong JW, Martinez AA, and Yan D. Defining a radiotherapy target with positron emission tomography. Int. J. Radiat. Oncol. Biol. Phys. 2004;60:1272-1282

17. Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M, and Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother. Oncol. 2003;69:247-250

18. van Dalen JA, Hoffmann AL, Dicken V, Vogel WV, Wiering B, Ruers TJ, Karssemeijer N, and Oyen WJ. A novel iterative method for lesion delineation and volumetric quantification with FDG-PET. Nucl. Med. Commun. (In press). 2007;

19. Schroeder W. The Visualization Toolkit: An object-oriented approach to 3D graphics, 3rd edn. New York: Kitware Inc. 2003;

20. Ibanez L, Schroeder W, Ng L, and Cates J. The ITK Software Guide: The Insight Segmentation and Registration Toolkit, version 1.4. New York: Kitware, Inc. 2003;

21. Kuhnigk JM, Dicken V, Bornemann L, Bakai A, Wormanns D, Krass S, and Peitgen HO. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. IEEE Trans. Med. Imaging. 2006;25:417-434

22. Drever L, Robinson DM, McEwan A, and Roa W. A local contrast based approach to threshold segmentation for PET target volume delineation. Med. Phys. 2006;33:1583-1594

23. Hashimoto Y, Tsujikawa T, Kondo C, Maki M, Momose M, Nagai A, Ohnuki T, Nishikawa T, and Kusakabe K. Accuracy of PET for diagnosis of solid pulmonary lesions with 18F-FDG uptake below the standardized uptake value of 2.5. J. Nucl. Med. 2006;47:426-431

24. Delbeke D, Martin WH, Sandler MP, Chapman WC, Wright JK, Jr., and Pinson CW. Evaluation of benign vs malignant hepatic lesions with positron emission tomography. Arch. Surg. 1998;133:510-515

25. Chin BB, Lyengar S, Sabundayo BP, and Schwartz D. Standardized Uptake Values in 2-Deoxy-2-[18F]Fluoro-D-Glucose with Positron Emission Tomography. Clinical Significance of Iterative Reconstruction and Segmented Attenuation Compared with Conventional Filtered Back Projection and Measured Attenuation Correction. Mol. Imaging Biol. 2002;4:294-300

26. Kaneta T, Hakamatsuka T, Takanami K, Yamada T, Takase K, Sato A, Higano S, Kinomura S, Fukuda H, Takahashi S, and Yamada S. Evaluation of the relationship between physiological FDG uptake in the heart and age, blood glucose level, fasting period, and hospitalization. Ann. Nucl. Med. 2006;20:203-208

27. Westerterp M, Pruim J, Oyen W, Hoekstra O, Paans A, Visser E, van LJ, Sloof G, and Boellaard R. Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. Eur. J. Nucl. Med. Mol. Imaging. 2006;

28. Sahani DV, Kalva SP, Fischman AJ, Kadavigere R, Blake M, Hahn PF, and Saini S. Detection of liver metastases from adenocarcinoma of the colon and pancreas: comparison of mangafodipir trisodium-enhanced liver MRI and whole-body FDG PET. AJR Am. J. Roentgenol. 2005;185:239-246

29. Scott DJ, Guthrie JA, Arnold P, Ward J, Atchley J, Wilson D, and Robinson PJ. Dual phase helical CT versus portal venous phase CT for the detection of colorectal liver metastases: correlation with intra-operative sonography, surgical and pathological findings. Clin. Radiol. 2001;56:235-242

30. Wiering B, Ruers TJ, Krabbe PF, Dekker HM, and Oyen WJ. Comparison of Multiphase CT, FDG-PET and Intra-Operative Ultrasound in Patients with Colorectal Liver Metastases Selected for Surgery. Ann. Surg. Oncol. 2007;14:818-826

31. Lee MJ, Saini S, Compton CC, and Malt RA. MR demonstration of edema adjacent to a liver metastasis: pathologic correlation. AJR Am. J. Roentgenol. 1991;157:499-501

32. Jentzen W, Freudenberg L, Eising EG, Heinze M, Brandau W, and Bockisch A. Segmentation of PET volumes by iterative image thresholding. J. Nucl. Med. 2007;48:108-114

33. Cohade C, Osman M, Leal J, and Wahl RL. Direct comparison of (18)F-FDG PET and PET/CT in patients with colorectal carcinoma. J. Nucl. Med. 2003;44:1797-1803

34. Vogel WV, Wiering B, Corstens FH, Ruers TJ, and Oyen WJ. Colorectal cancer: the role of PET/CT in recurrence. Cancer Imaging. 2005;5 Spec No A:S143-S149

35. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, and Corstens FH. PET/CT: panacea, redundancy, or something in between? J. Nucl. Med. 2004;45 Suppl 1:15S-24S

14

Discussion and future prospects

14.1 Advances in PET imaging

Although the first positron emission tomography (PET) scanner has been built as early as in 1953, improvements are constantly being achieved in e.g. image quality, determination of diagnostic value, and in the selection of indications (1).

Image quality

Spatial resolution, contrast resolution, and noise levels are relevant parameters in the diagnostic performance of PET imaging. The image quality is influenced by many factors, such as scanner resolution and sensitivity, positron range (2), injected activity (3), biodistribution time, patient characteristics, and image reconstruction. All these factors require specific optimizations.

Multiple improvements in scanner design are to be expected in the coming decade, with significantly better image quality as a result. Examples are more efficient detection of photons, incorporation of annihilation photon flight time information (4,5), and adaptation of scanner design to specific clinical requirements (6-8). However, these advances are not the only issues that require attention. Several of the abovementioned factors in image quality are not scanner-related. Therefore, additional efforts need to be directed at patient-related factors, such as weight-adapted activity administration, adapted acquisition time per bed position, adapted biodistribution times, and optimized image reconstruction. An example of the latter is the improvement of the diagnostic yield of FDG-PET through adaptation of image reconstruction parameters to low photon-attenuation in the head and neck area (chapter 3), or other specific body characteristics (e.g. children or adipose patients). Incorporation of such factors may lead to a more variable approach to PET imaging, tailored to a specific patient and a particular disease. As a consequence, advances beyond the current imaging quality will require a more active role of individual nuclear medicine departments.

Radiation dose

An additional benefit of improved scanner efficiency lies in dosimetric considerations. In Europe, traditional protocols require an injected activity in the range of 200–400 MBq FDG with an average of 370 MBq (3). Such doses allow wholebody emission acquisition in 20-30 minutes while delivering an acceptable effective radiation dose of 4-8 mSv to the patient. Recent technical advances have already allowed the same image quality while reducing either the injected activity or the acquisition time by half, or with a balanced benefit of both factors (5). Optimal PET imaging involves a balance between injected activity, patient throughput per (expensive) scanner, and achieved image quality. In the current economical timeframe, high thoughput is a necessity in many diagnostic centers. From the patient's perspective, it is to be hoped that better availability of PET imaging will soon allow a more prudent throughput, resulting in a lower radiation dose while refraining from compromises to image quality.

Standardization of imaging

The abovementioned considerations apply according to local factors. This has led to differences between neighbouring imaging centers, in image quality, diagnostic value, and applied radiation dose. This may result in unacceptable impact on e.g. SUV quantification (9) or clinical value, and underlines the importance of standardized imaging protocols. Current efforts are focussed on reduction of inter-hospital differences in acquisition and quantification of PET analysis (9). As a subsequent step, minimal requirements could be formulated with regard to scanner sensitivity, image quality, and delivered radiation dose.

Tracer development

A further benefit may be achieved with new radiopharmaceuticals. FDG has been the most widely used tracer for years, because of a favorably high lesion-to-background ratio for many disorders, and thus a very high sensitivity. Nevertheless, FDG does have drawbacks. Accumulation in both malignant and benign (inflammatory) diseases limits specificity. Many tracers are being developed that promise to be more specific, such as ^{18}F-FLT for imaging of DNA synthesis in primary tumors and its metastases (10). These new radiopharmaceuticals need proper validation prior to clinical implementations, because not all expectations may be fulfilled (i.e. staging of lymph nodes with ^{18}F-FLT in cancer of the head and neck area, chapter 8). Other promising examples currently under investigation include ^{18}F-FET for specific tumor imaging by visualization of protein synthesis (11-13), and ^{11}C-acetate for imaging of non FDG-avid malignancies (14,15).

Theoretically, short-lived radionuclides may be preferred to ^{18}Fluor. These nuclides often deliver a relatively low radiation dose and allow multiple acquisitions per day, although this will require the introduction of on-site cyclotrons or generators as standard facilities. Much work still needs to be done to expand the spectrum of clinically useful tracers.

Indications

As PET imaging requires medical resources and utilizes ionizing radiation, clinical indications need to be carefully selected. Important criteria are measurable benefits for the patient, e.g. fewer complaints and complications, less futile invasive procedures, a longer disease-free interval, or ultimately a longer survival due to better selection or optimization of treatment. Research has resulted in a clear benefit of FDG-PET, as illustrated by current protocols in the Netherlands, for e.g. diagnosis and staging of lung cancer (16), metastatic colorectal cancer (17), and lymphoma. When PET does not provide sufficient clinical benefit, it should not be incorporated in clinical practice, as was demonstrated for staging of clinically node-negative head and neck cancer (chapter 4).

14.2 *Molecular imaging*

The concept of molecular imaging is not limited to nuclear medicine, as selected molecules can be visualized in vivo with other techniques as well (18). Substitution of the radioactive label with a fluorescent label allows a similar approach using optical fluorescence imaging, while a paramagnetic label allows imaging using magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS) (19).

New applications of molecular imaging are suggested frequently. Recent examples include exciting new fields such as monitoring of gene therapy (20,21), imaging of stem cell differentiation (22), and adaptation of external beam radiation therapy to various local tissue characteristics (23,24). Molecular imaging modalities each have their characteristics and require validation and positioning in clinical practice, either alone or in an integrated approach.

Molecular imaging modalities

Optical imaging, with its very high spatial and temporal resolution, is limited to analysis of superficial tissues due to the physical characteristics of (near infrared) visible light (25). This restricts applications to small animal imaging, ex vivo imaging of tissues, and superficial lesions. In the currents implementation of optical imaging, the acquisition of a quantitative signal is still a challenge (26).

(Functional) MRI provides in vivo imaging throughout the body with high spatial and temporal resolution (27,28). Currently, the number of available paramagnetic molecules is still limited, but expanding. Downscaling of the relatively high paramagnetic tracer dose remains a challenge. Furthermore, the technique is usually limited to a region of interest within the body due to the relatively long image acquisition times, and exact quantification is still an issue (29).

PET imaging provides whole-body and quantitative biodistribution imaging in both small animals (30) and humans (31). The excellent sensitivity allows imaging of many biologically active molecules in physiological quantities. On the other hand, spatial resolution is relatively low and the technique inherently results in a radiation burden to the patient and, to a lesser extent, to the operators.

The abovementioned characteristics overlap in many aspects for the various imaging techniques. As the chemical procedures for labeling and the approaches to interpretation are comparable for the respective imaging modalities, collaboration in research is needed. Many disciplines, such as pathology, radiology, radiation oncology, and nuclear medicine use implementations of molecular imaging for their research. Coordinated use of knowledge and equipment can provide an effective integrated research setting in terms of productivity and costs. This approach requires logistical cooperation, rather than technological advances. The success of an integrated approach is currently illustrated by the good results of emerging multidisciplinary *imaging centers*, where education, research and clinical diagnostics surpass departmental borders.

14.3 Advances in multimodality imaging

Developing technology and constantly increasing evidence, as described for PET imaging, are unabatedly applicable to multimodality imaging. It is only just becoming clear in which cases molecular imaging with PET has clinical value as a standalone modality, or in addition to anatomical imaging or other functional imaging modalities. And when both functional and anatomical imaging are available, it is an unresolved issue in what way the images may best be correlated or combined (chapters 2 and 9).

Nevertheless, technical advances have already allowed the development of integrated PET/CT scanning devices (32). Some scientific evidence of the added value in general application of these devices has been published (33,34), as well as support for some specific clinical indications. However, at present there is insufficient evidence to classify the supposed supremacy of hybrid PET/CT over dedicated imaging as "evidence based medicine". Nevertheless, these combined devices have found their way into many hospitals, although several important issues have not yet been resolved (chapters 2 and 11).

Clear strong points of integrated PET/CT scanning – convincing although difficult to prove – are the logistical benefit and the improved understanding of fused images. Many groups have demonstrated the additional value of integrated imaging on clinical decision making and therapy planning (34-38). Based on these considerations, it is to be expected that integrated PET/CT scanners will prevail.

Given this assumption, it is important that the drawbacks of this technique are acknowledged (i.e. imperfect image registration and attenuation correction artefacts, chapter 2), to allow proper selection of protocols and adaptation of reviewing (chapter 11). With these precautions, integrated PET/CT in its current state may be considered a safe and adequate solution until better alternatives will be developed. A first and important improvement could be expected from re-introduction of "classical" transmission imaging for attenuation correction in hybrid PET/CT (chapter 11), possibly synchronously acquired during emission imaging (39). A second strategy that deserves further evaluation is breathing-gated acquisition of both PET and CT, to eliminate motion and positional differences between the image sets (40,41).

SPECT/CT

Many of these issues apply to integrated SPECT/CT scanning as well (42), although to a lesser extent. The lower spatial resolution of SPECT studies will bring image correlation to a coarser and less critical scale. Attenuation correction can be applied using CT images with less risk for artefacts, because the applied gamma photons tend to have photon energies closer to X-ray photons. SPECT images may benefit even more from correlation with anatomical imaging than many PET studies (43). Similar to integrated PET/CT, the combination of SPECT and CT seems to benefit interpretation and understanding of the images (44-47), and is thus likely to gain popularity. Nevertheless, SPECT/CT is a new technique with specific issues, that need proper validation (48).

PET/MRI

Molecular imaging with PET generally involves evaluation of (processes within) soft tissues, for which correlation with CT may not be the optimal choice (chapter 9). Many soft tissue processes can be localized and delineated with better contrast and accuracy using MRI. Therefore, a combination of PET with MRI may be preferable. As a secondary benefit, MRI does not contribute to the radiation dose, as does CT. The combination of PET and MRI in one single device is a major technical challenge. However, a working proof of concept has already been demonstrated (49), and clinical scanners may be expected within the next 5 years.

Software fusion

Where integrated imaging is not available, and for combinations of scans other than PET/CT or SPECT/CT, software fusion of dedicated image sets remains an attractive alternative. As described in this thesis, software fusion is adequate in organ-focused approaches. Examples are the head and neck area when rigidly fixed (chapter 6), and the liver when isolated from surrounding organs (chapter 10). Intuitively, other organ-focused methods will perform similarly, when tailored solutions for specific problems are addressed. Based on our experiences and the lack of validated tools, unrestrained image fusion of whole-body studies must be discouraged, or should be interpreted with sufficient reserve. This also applies to upcoming non-rigid (elastic) software fusion. Algorithm-based 3-dimensional deformation of image sets is unlikely to behave identical to a human body, including unpredictable phenomena like parenchymal deformation and pleural displacement. Corruption of images and subsequent misinterpretation need to be ruled out by supporting evidence, prior to clinical implementation.

In summary, integration of images can be useful and will continue to expand its field of applications, but is yet to find its final place and implementations.

14.4 Localization and delineation of disease

Exact anatomical localization and delineation of disease is essential for both diagnostic procedures and planning of therapeutic interventions. Many imaging modalities can contribute to this clinical challenge, but no single modality can delineate any type of lesion with 100% accuracy. Therefore, the individual characteristics of imaging techniques need to be weighed, and may be combined as a logical next step.
Despite its superior contrast resolution, the place of FDG-PET imaging in the spectrum of available modalities is not fully elucidated. A relatively poor spatial resolution and “blurry” appearance of lesions demand a tailored approach to lesion delineation, of which many different implementations currently exist, each with their own characteristics and applications. Examples are visual analysis, contouring based on lesion or background signal intensities, and adaptive thresholding techniques based on different variables. The choice of strategy depends on locoregional tissue and scanner characteristics. For example, blurring, noise and diffuse background activity have a strong influence in PET imaging of the liver (chapters 12 and 13), while variable uptake in surrounding normal tissues is the most important limiting factor for delineation of tumor in the head and neck area (chapter 7). Several advances have been described in this thesis, but many questions remain, and lesion delineation with PET is still an issue. Further steps are necessary, and will be easier to take when images with higher spatial resolution become available.

As in any clinical application, for lesion delineation the advantage of PET (sensitive lesion detection) must be weighed against advantages of other imaging modalities, such as ultrasound, CT, and MRI (good spatial resolution and anatomical orientation). For the head and neck area, reliable PET/CT image registration with an accuracy in a clinically acceptable range could be verified in this thesis (chapters 5 and 6), with subsequent successful application in external beam radiation treatment target volume definition (chapter 7). However, in the upper abdominal area, integrated PET/CT imaging in its current state cannot guarantee perfect image registration (chapters 2 and 11). As shown in this thesis (chapter 10), specific software fusion for multimodality imaging of the liver can be highly optimized, but is laborious compared to hybrid PET/CT imaging. Lesion delineation in the liver was validated for both PET and CT (chapter 13), but the procedure as a whole will further benefit from new optimizations in image registration.

These experiences once again suggest that lesion delineation requires a tailored approach, optimized to locoregional tissue characteristics and multimodality imaging possibilities. Much work is to be done in this area of diagnostic imaging.

14.5 Standardization in medical imaging

The purpose of modern medicine is *proven benefit* for patients in prevention, curation, and supportive care. Determination of such benefit is complicated in current medical imaging, due to overlapping indications for the various anatomical and functional imaging techniques. This has resulted in implementations adapted to local possibilities and preferences. Selected imaging modalities are applied with different (and ever improving) parameters that are locally available and that may influence clinical value (e.g. CT scanners with an increasing number of parallel detector rows, MRI scanners with increasing field strength and novel imaging sequences, and a variety of contrast enhancing substances). Furthermore, within a single imaging modality, imaging devices from different manufacturers may vary in specifications. The spectrum is further extended by innovative visual representation methods such as 3-dimensional reconstructions, computer aided diagnosis, and image fusion. Such advances are promising and find rapid acceptance in clinical strategies. But at this point it is difficult to determine the relative value of different approaches, with regard to patient benefit or cost effectiveness. These factors result in highly variable implementations and interpretation of diagnostic strategies across hospitals (50). It has become a challenge to determine which diagnostic tests are most appropriate for a specific entity or disease.

The rapidly increasing number – and high turnover – of available diagnostic strategies also complicates the setup and interpretation of multi-center trials (50). The majority of recent scientific publications on diagnostic imaging comprise single-center studies, as exemplified by this thesis. Such research must be interpreted in the perspective of local opportunities and experience.

In the forthcoming years, some of the abovementioned issues will be resolved. At the same time, however, new issues are likely to arise from novel diagnostic techniques, such as e.g. new PET tracers and integrated PET/MRI. From the patient’s perspective, a critical attitude towards new technological achievements remains important.

14.6 Cancer in the head and neck area

The final positioning of PET in imaging of cancer in the head and neck area still requires an effort. Few generally accepted protocols exist for either primary diagnostic imaging, staging, relapse detection, re-staging, therapy monitoring, or therapy planning with FDG or any other PET radiopharmaceutical.

The protocols that do exist all apply to FDG, for imaging and treatment monitoring of primary tumors. However, FDG-PET is not an ideal technique for imaging of malignancy in the head and neck area, due to visualization of concurrent local inflammation, especially during and shortly after treatment. FLT and other more specific tracers than FDG hold much promise for better delineation and characterization of tumors (51).

For initial lymph node staging, FDG-PET imaging cannot contribute to standard imaging protocols (chapter 4) despite extensive optimizations (chapter 3), and FLT is not the solution either due to confounding B-cell proliferation (chapter 8). Nevertheless, several new approaches have already been suggested, e.g. staging with acetate (52), treatment monitoring with FLT (53), integration of PET with diagnostic CT (54), and correlation with MRI (55). The clinical value of these new approaches is yet to be determined.

14.7 Liver metastases from colorectal cancer

In parallel, imaging of liver metastases remains challenging. Lesions may be very small, irregularly shaped, isodense with surrounding normal liver tissue, and with limited changes in perfusion. This poses problems for each of the applied imaging modalities.

Lesion delineation with FDG-PET can now be performed with uncertainties in a similar range as with CT imaging, as shown in this thesis (chapters 12 and 13). The number of studies that suggest a clinically significant place for FDG-PET imaging of liver metastasis is increasing (56,57). However, FDG is not the ideal PET tracer for imaging of liver metastases, due to uptake in normal liver tissue. As a consequence, adequate delineation of small liver metastases is difficult (57). Future improvements in scanner efficiency and resolution will contribute positively to this issue, but will not resolve it entirely. New tracers with improved lesion-to-background ratios are awaited, but not expected within the foreseeable future.

The combination of FDG-PET and (contrast-enhanced) CT for imaging of liver metastases is valuable (58-60) (chapters 9 and 13), but not perfect (chapters 10 and 11). Breathing-related misalignment and artefact issues play a prominent role in the area of the diaphragmatic dome in current integrated PET/CT imaging. For lesion delineation and correlative localization with anatomical structures, artefact-free well-correlated images are crucial. These requirements may be achieved through re-implementation of “classical” transmission imaging in integrated PET/CT, and/or gated acquisition of both PET and CT images. While such optimizations are awaited, registered image sets need to be interpreted with caution. Meanwhile, software fusion of dedicated PET and CT can provide reliable correlation on an incidental basis (chapter 10).

14.8 References

1. Cherry SR. The 2006 Henry N. Wagner Lecture: Of mice and men (and positrons) - advances in PET imaging technology. J. Nucl. Med. 2006;47:1735-1745
2. Pagani M, Stone-Elander S, and Larsson SA. Alternative positron emission tomography with non-conventional positron emitters: effects of their physical properties on image quality and potential clinical applications. Eur. J. Nucl. Med. 1997;24:1301-1327
3. Brix G, Nosske D, Glatting G, Minkov V, and Reske SN. A survey of PET activity in Germany during 1999. Eur. J. Nucl. Med. Mol. Imaging. 2002;29:1091-1097
4. Mullani NA, Markham J, and Ter-Pogossian MM. Feasibility of time-of-flight reconstruction in positron emission tomography. J. Nucl. Med. 1980;21:1095-1097
5. Surti S, Karp JS, Popescu LM, ube-Witherspoon ME, and Werner M. Investigation of time-of-flight benefit for fully 3-D PET. IEEE Trans. Med. Imaging. 2006;25:529-538
6. Abreu MC, Almeida P, Balau F, Ferreira NC, Fetal S, Fraga F, Martins M, Matela N, Moura R, Ortigao C, Peralta L, Rato P, Ribeiro R, Rodrigues P, Santos AI, Trindade A, and Varela J. Clear-PEM: a dedicated PET camera for improved breast cancer detection. Radiat. Prot. Dosimetry. 2005;116:208-210
7. Braem A, Llatas MC, Chesi E, Correia JG, Garibaldi F, Joram C, Mathot S, Nappi E, da Silva MR, Schoenahl F, Seguinot J, Weilhammer P, and Zaidi H. Feasibility of a novel design of high resolution parallax-free Compton enhanced PET scanner dedicated to brain research. Phys. Med. Biol. 2004;49:2547-2562
8. Qi J. Optimization of a PET scanner design for prostate lesion detection. Conf. Proc. IEEE Eng Med. Biol. Soc. 2004;2:1357-1360
9. Westerterp M, Pruim J, Oyen W, Hoekstra O, Paans A, Visser E, van LJ, Sloof G, and Boellaard R. Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. Eur. J. Nucl. Med. Mol. Imaging. 2006;
10. Been LB, Suurmeijer AJ, Cobben DC, Jager PL, Hoekstra HJ, and Elsinga PH. [18F]FLT-PET in oncology: current status and opportunities. Eur. J. Nucl. Med. Mol. Imaging. 2004;31:1659-1672
11. Chang CH, Wang HE, Wu SY, Fan KH, Tsai TH, Lee TW, Chang SR, Liu RS, Chen CF, Chen CH, and Fu YK. Comparative evaluation of FET and FDG for differentiating lung carcinoma from inflammation in mice. Anticancer Res. 2006;26:917-925

12. Floeth FW, Pauleit D, Sabel M, Reifenberger G, Stoffels G, Stummer W, Rommel F, Hamacher K, and Langen KJ. 18F-FET PET differentiation of ring-enhancing brain lesions. J. Nucl. Med. 2006;47:776-782

13. Laique S, Egrise D, Monclus M, Schmitz F, Garcia C, Lemaire C, Luxen A, and Goldman S. L-amino acid load to enhance PET differentiation between tumor and inflammation: an in vitro study on (18)F-FET uptake. Contrast. Media Mol. Imaging. 2006;1:212-220

14. Liu RS, Chang CP, Chu LS, Chu YK, Hsieh HJ, Chang CW, Yang BH, Yen SH, Huang MC, Liao SQ, and Yeh SH. PET imaging of brain astrocytoma with 1-11C-acetate. Eur. J. Nucl. Med. Mol. Imaging. 2006;33:420-427

15. Oyama N, Miller TR, Dehdashti F, Siegel BA, Fischer KC, Michalski JM, Kibel AS, Andriole GL, Picus J, and Welch MJ. 11C-acetate PET imaging of prostate cancer: detection of recurrent disease at PSA relapse. J. Nucl. Med. 2003;44:549-555

16. Richtlijn niet-kleincellig longcarcinoom: stadiëring en behandeling. Vereniging van Integrale Kankercentra, Kwaliteitsinstituut voor de Gezondheidszorg CBO, the Netherlands.

17. Wiering B, Krabbe PF, Jager GJ, Oyen WJ, and Ruers TJ. The impact of fluor-18-deoxyglucose-positron emission tomography in the management of colorectal liver metastases. Cancer. 2005;104:2658-2670

18. Weissleder R. Molecular imaging in cancer. Science. 2006;312:1168-1171

19. Shah GV, Fischbein NJ, Patel R, and Mukherji SK. Newer MR imaging techniques for head and neck. Magn Reson. Imaging Clin. N. Am. 2003;11:449-69, vi

20. Iyer M, Sato M, Johnson M, Gambhir SS, and Wu L. Applications of molecular imaging in cancer gene therapy. Curr. Gene Ther. 2005;5:607-618

21. Nichol C and Kim EE. Molecular imaging and gene therapy. J. Nucl. Med. 2001;42:1368-1374

22. Sheikh AY and Wu JC. Molecular imaging of cardiac stem cell transplantation. Curr. Cardiol. Rep. 2006;8:147-154

23. Brahme A. Biologically optimized 3-dimensional in vivo predictive assay-based radiation therapy using positron emission tomography-computerized tomography imaging. Acta Oncol. 2003;42:123-136

24. Grosu AL, Piert M, Weber WA, Jeremic B, Picchio M, Schratzenstaller U, Zimmermann FB, Schwaiger M, and Molls M. Positron emission tomography for radiation treatment planning. Strahlenther. Onkol. 2005;181:483-499

25. Ntziachristos V, Bremer C, and Weissleder R. Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging. Eur. Radiol. 2003;13:195-208

26. Gandjbakhche AH, Chernomordik V, Hattery D, Hassan M, and Gannot I. Tissue characterization by quantitative optical imaging methods. Technol. Cancer Res. Treat. 2003;2:537-551

27. Jasanoff A. Functional MRI using molecular imaging agents. Trends Neurosci. 2005;28:120-126

28. Winter PM, Caruthers SD, Wickline SA, and Lanza GM. Molecular imaging by MRI. Curr. Cardiol. Rep. 2006;8:65-69

29. Morawski AM, Winter PM, Crowder KC, Caruthers SD, Fuhrhop RW, Scott MJ, Robertson JD, Abendschein DR, Lanza GM, and Wickline SA. Targeted nanoparticles for quantitative imaging of sparse molecular epitopes with MRI. Magn Reson. Med. 2004;51:480-486

30. Chatziioannou AF. Molecular imaging of small animals with dedicated PET tomographs. Eur. J. Nucl. Med. Mol. Imaging. 2002;29:98-114

31. Price P. PET as a potential tool for imaging molecular mechanisms of oncology in man. Trends Mol. Med. 2001;7:442-446

32. von Schulthess GK, Steinert HC, and Hany TF. Integrated PET/CT: current applications and future directions. Radiology. 2006;238:405-422

33. Czernin J, len-Auerbach M, and Schelbert HR. Improvements in Cancer Staging with PET/CT: Literature-Based Evidence as of September 2006. J. Nucl. Med. 2007;48 Suppl 1:78S-88S

34. Sachelarie I, Kerr K, Ghesani M, and Blum RH. Integrated PET-CT: evidence-based review of oncology indications. Oncology (Williston. Park). 2005;19:481-490

35. Blodgett TM, Meltzer CC, and Townsend DW. PET/CT: form and function. Radiology. 2007;242:360-385

36. Ell PJ. PET/CT in oncology: a major technology for cancer care. Chang Gung. Med. J. 2005;28:274-283

37. von Schulthess GK. Positron emission tomography versus positron emission tomography/computed tomography: from “unclear” to “new-clear” medicine. Mol. Imaging Biol. 2004;6:183-187

38. Wechalekar K, Sharma B, and Cook G. PET/CT in oncology--a major advance. Clin. Radiol. 2005;60:1143-1155

39. Matsumoto K, Kitamura K, Mizuta T, Tanaka K, Yamamoto S, Sakamoto S, Nakamoto Y, Amano M, Murase K, and Senda M. Performance characteristics of a new 3-dimensional continuous-emission and spiral-transmission high-sensitivity and high-resolution PET camera evaluated with the NEMA NU 2-2001 standard. J. Nucl. Med. 2006;47:83-90

40. Nehmeh SA, Erdi YE, Pan T, Pevsner A, Rosenzweig KE, Yorke E, Mageras GS, Schoder H, Vernon P, Squire O, Mostafavi H, Larson SM, and Humm JL. Four-dimensional (4D) PET/CT imaging of the thorax. Med. Phys. 2004;31:3179-3186

41. Wolthaus JW, van HM, Muller SH, Belderbos JS, Lebesque JV, de Bois JA, Rossi MM, and Damen EM. Fusion of respiration-correlated PET and CT scans: correlated lung tumour motion in anatomical and functional scans. Phys. Med. Biol. 2005;50:1569-1583

42. Schillaci O. Hybrid SPECT/CT: a new era for SPECT imaging? Eur. J. Nucl. Med. Mol. Imaging. 2005;32:521-524

43. Schillaci O, Danieli R, Manni C, and Simonetti G. Is SPECT/CT with a hybrid camera useful to improve scintigraphic imaging interpretation? Nucl. Med. Commun. 2004;25:705-710

44. Filippi L and Schillaci O. SPECT/CT with a hybrid camera: a new imaging modality for the functional anatomical mapping of infections. Expert. Rev. Med. Devices. 2006;3:699-703

45. Ingui CJ, Shah NP, and Oates ME. Endocrine neoplasm scintigraphy: added value of fusing SPECT/CT images compared with traditional side-by-side analysis. Clin. Nucl. Med. 2006;31:665-672

46. Lerman H, Metser U, Lievshitz G, Sperber F, Shneebaum S, and Even-Sapir E. Lymphoscintigraphic sentinel node identification in patients with breast cancer: the role of SPECT-CT. Eur. J. Nucl. Med. Mol. Imaging. 2006;33:329-337

47. Savelli G, Maffioli L, Maccauro M, De DE, and Bombardieri E. Bone scintigraphy and the added value of SPECT (single photon emission tomography) in detecting skeletal lesions. Q. J. Nucl. Med. 2001;45:27-37

48. Kuikka JT. Myocardial perfusion imaging with a novel SPECT/CT system: all that glitters is not gold. Eur. J. Nucl. Med. Mol. Imaging. 2007;

49. Pichler BJ, Judenhofer MS, Catana C, Walton JH, Kneilling M, Nutt RE, Siegel SB, Claussen CD, and Cherry SR. Performance test of an LSO-APD detector in a 7-T MRI scanner for simultaneous PET/MRI. J. Nucl. Med. 2006;47:639-647

50. Shankar LK, Hoffman JM, Bacharach S, Graham MM, Karp J, Lammertsma AA, Larson S, Mankoff DA, Siegel BA, Van den AA, Yap J, and Sullivan D. Consensus recommendations for the use of 18F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. J. Nucl. Med. 2006;47:1059-1066

51. Cobben DC, van der Laan BF, Maas B, Vaalburg W, Suurmeijer AJ, Hoekstra HJ, Jager PL, and Elsinga PH. 18F-FLT PET for visualization of laryngeal cancer: comparison with 18F-FDG PET. J. Nucl. Med. 2004;45:226-231

52. Sun A, Sorensen J, Karlsson M, Turesson I, Langstrom B, Nilsson P, Cederblad L, Bertling J, Riklund K, and Johansson S. 1-[(11)C]-acetate PET imaging in head and neck cancer-a comparison with (18)F-FDG-PET: implications for staging and radiotherapy planning. Eur. J. Nucl. Med. Mol. Imaging. 2006;

53. Barthel H, Cleij MC, Collingridge DR, Hutchinson OC, Osman S, He Q, Luthra SK, Brady F, Price PM, and Aboagye EO. 3’-deoxy-3’-[18F]fluorothymidine as a new marker for monitoring tumor response to antiproliferative therapy in vivo with positron emission tomography. Cancer Res. 2003;63:3791-3798

54. Jeong HS, Baek CH, Son YI, Ki CM, Kyung LD, Young CJ, Kim BT, and Kim HJ. Use of integrated (18)F-FDG PET/CT to improve the accuracy of initial cervical nodal evaluation in patients with head and neck squamous cell carcinoma. Head Neck. 2006;

55. Daisne JF, Sibomana M, Bol A, Cosnard G, Lonneux M, and Gregoire V. Evaluation of a multimodality image (CT, MRI and PET) coregistration procedure on phantom and head and neck cancer patients: accuracy, reproducibility and consistency. Radiother. Oncol. 2003;69:237-245

56. Bipat S, van Leeuwen MS, Comans EF, Pijl ME, Bossuyt PM, Zwinderman AH, and Stoker J. Colorectal liver metastases: CT, MR imaging, and PET for diagnosis--meta-analysis. Radiology. 2005;237:123-131

57. Wiering B, Krabbe PF, Dekker HM, Oyen WJ, and Ruers TJ. The Role of FDG-PET in the Selection of Patients with Colorectal Liver Metastases. Ann. Surg. Oncol. 2007;14:771-779

58. Erturk SM, Ichikawa T, Fujii H, Yasuda S, and Ros PR. PET imaging for evaluation of metastatic colorectal cancer of the liver. Eur. J. Radiol. 2006;58:229-235

59. Khan S, Tan YM, John A, Isaac J, Singhvi S, Guest P, and Mirza DF. An audit of fusion CT-PET in the management of colorectal liver metastases. Eur. J. Surg. Oncol. 2006;32:564-567

60. Selzner M, Hany TF, Wildbrett P, McCormack L, Kadry Z, and Clavien PA. Does the novel PET/CT imaging modality impact on the treatment of patients with metastatic colorectal cancer of the liver? Ann. Surg. 2004;240:1027-1034

15

Summary and conclusions

15.1 Summary

Several advances in functional and multimodality imaging were discussed in this thesis. The characteristics of PET and combined PET/CT imaging were reviewed (chapter 2), and subsequently, multiple hypotheses and possibilities regarding improved imaging were tested, with implications for clinical PET imaging.

Diagnosis and therapy of cancer in the head and neck area

The image quality of PET was optimized for detection of cancer in the head and neck area (chapter 3), resulting in better lymph node staging and improved management of patients with cancer of the head and neck area. Since routine whole-body PET reconstruction parameters proved to be inadequate for the head and neck area, and optimized parameters are scanner-specific, it is advised to implement optimization procedures for all PET scanners. Despite these optimizations, in patients with cancer in the head and neck area and without clinical signs of lymph node metastases, it was concluded that PET imaging could not replace invasive supra-omohyoidal lymph node dissection (chapter 4).

Multimodality imaging with PET and CT was optimized and validated for application in the head and neck area. After correction of an unexpected difference in real image size between PET and CT (chapter 5), and after selection of the most appropriate approach to image registration, software image registration could be validated for implementation in novel 3-dimensional intensity modulated radiotherapy (chapter 6). The spatial accuracy in image registration that could be achieved when using the iterative closest point algorithm (better than 3 mm in the planning area) now permits the application of software image fusion for IMRT where dedicated PET and CT are available, with no need for laborious fiducial markers.

Subsequent analysis of lesion delineation on FDG-PET images, for implementation in IMRT field planning, demonstrated that the clinical relevance of this approach is potentially high, but on the other hand that additional work needs to be performed to evaluate the clinical impact of the introduction of FDG-PET in radiation treatment planning (chapter 7). Multiple delineation methods for PET are available and in use, but the choice of methods has a major impact on the resulting gross target volume (GTV). These variations, which exist also relative to CT-based planning, have led to the conclusion that PET-based GTV definition in the head and neck area as yet may not be considered a validated technique.

The detection of lymph node metastases in the neck with FDG-PET is hampered by a low specificity, due to visualization of inflammation. FLT, which accumulates in proliferating cells, was hypothesized to have a higher specificity than FDG. However, false positive visualization of inflammatory lymph nodes also occurred with FLT, due to accumulation in proliferating B-cells (chapter 8). It was concluded that FLT-PET cannot replace FDG-PET in lymph node staging in patients with cancer in the head and neck area, and that the ideal tracer for PET imaging for this specific application is yet to be discovered.

Imaging of liver metastasis in colorectal cancer

The combination of PET and CT is currently proving itself as a valuable tool in the diagnostic strategy for detection of recurrent colorectal carcinoma, especially in the field of staging before surgical reinterventions, with an impact on diagnosis and choice of therapy. In this view, the applications of separate PET and CT, and of integrated PET/CT, were reviewed (chapter 9).

The performance of software image fusion of PET and CT of the liver was improved by the introduction of a new organ-focussed and observer-independent image registration method (chapter 10). The procedure demonstrated a precision of 2.5 ± 1.3 mm throughout the liver, and was thus validated for clinical applications, although the procedure was found to be laborious compared to co-registration using hybrid PET/CT, and required specific software and skills.

Subsequently, the performance of integrated PET/CT for imaging of the liver was evaluated (chapter 11). Anatomical registration errors of the liver in PET/CT proved to be potentially severe, occurring mainly due to breathing differences during acquisition of PET and CT. These errors resulted in clinically relevant correlation errors at the location of the diaphragm in 40-55% of cases, depending on the applied breathing protocol. Subsequent significant attenuation correction artefacts could occur where registration errors and a sharp transition between dense and non-dense tissue co-existed. It was determined that the magnitude of image registration errors and attenuation correction errors cannot be appreciated visually on PET images that have been corrected for photon attenuation. Based on these findings, recommendations were formulated for optimal imaging and reviewing of integrated PET/CT.

For improved delineation of liver metastases on PET, a background-subtracted relative threshold level method was designed (chapter 12). This method proved to be independent of the signal-to-background ratio and the size of the lesions, and was validated in phantom and in patient studies. Subsequently, the performance of this delineation method for size measurements of liver metastases was determined, in comparison with other available delineation methods for PET and CT (chapter 13). By correlation with pathological or intra-operative verification of measured lesion diameters, the accuracy of the measurement of liver metastases on FDG-PET images could be validated for clinical application.

15.2 Conclusions

In this thesis, several advances in PET and PET/CT imaging have been achieved. Imaging and processing procedures have been optimized, and clinical applications have been validated. Some issues still remain unclear or unresolved. The spatial resolution of molecular imaging with PET will remain relatively low as compared to CT and MRI, and combined PET/CT imaging will bear risks of correlative errors and artefacts in its current implementations. Localization and delineation of tumor tissue with PET remains a good approximation, whether performed with stand-alone imaging or with integrated modalities.

PET imaging and multimodality imaging will require further optimization. Meanwhile, awareness and understanding of the potential and remaining issues can help in achieving adequate interpretation of the images, and can guide selection of additional clinical indications. Further implementation of PET imaging and multimodality imaging in consensus protocols is mandatory, in order to reduce regional differences in applications and clinical value, and to allow comparative multi-center investigations. The continuation of such research is essential and must synchronize with the current rapidly evolving technical advances.

16

Samenvatting in het Nederlands

16.1 Samenvatting

Positron emissie tomografie (PET) is een relatief nieuwe techniek waarmee radioactieve stoffen binnen in het menselijk lichaam kunnen worden afgebeeld, gevolgd en gekwantificeerd. Met de meestgebruikte radioactieve stof, fluor-18 gekoppeld aan een variant van suiker (FDG), kan het energieverbruik (metabolisme) van de verschillende organen en weefsels binnen een patiënt in beeld worden gebracht. Met FDG-PET kunnen belangrijke vragen worden beantwoord op het gebied van de diagnostiek van kwaadaardige ziekten. Het afbeelden van de activiteiten en eigenschappen van weefsels wordt 'functionele beeldvorming' genoemd.

In veel gevallen is voor het stellen van een juiste diagnose ook informatie nodig over de exacte locatie en het formaat van zieke weefsels, en de relatie met omliggende gezonde organen. Deze gegevens worden verkregen met 'anatomische beeldvorming', zoals computer tomografie (CT) en magnetische resonantie imaging (MRI).

In de klinische praktijk zijn vaak zowel anatomische als functionele beelden van een patiënt beschikbaar, en kan de meest nauwkeurige diagnose worden gesteld door onderlinge vergelijking van alle resultaten. Daarvoor kunnen verschillende beelden naast elkaar, maar ook óver elkaar worden geprojecteerd. Dit laatste staat bekend als beeldfusie. Vanwege de vaak waardevolle synergistische informatie in deze gecombineerde beelden zijn in de afgelopen jaren gecombineerde apparaten op de markt gebracht, die verschillende typen scans in één sessie kunnen maken.

Een voorbeeld daarvan is de gecombineerde PET/CT-scanner. Een dergelijke combinatie van verschillende apparaten heeft belangrijke voordelen, zoals snellere logistiek, handige vergelijking van beelden, en soms betere uitslagen. Maar er kunnen ook nadelen zijn. Zo is een gecombineerd apparaat duurder en is de stralingsbelasting voor patienten soms hoger, terwijl de gecombineerde beelden niet altijd nodig zijn. In bepaalde gevallen kunnen de verschillende componenten elkaars beeldkwaliteit nadelig beinvloeden. De uiteindelijke plaats van gecombineerde scanners in de medische diagnostiek is daarom nog onderwerp van onderzoek.

In dit proefschrift werden meerdere ontwikkelingen in de functionele beeldvorming met PET belicht. De karakteristieken van PET en gecombineerde PET/CT werden besproken (hoofdstuk 2). Vervolgens werden verschillende hypothesen en technieken met betrekking tot verbeterde beeldvorming getest. Hieruit werden aanbevelingen afgeleid voor verbeterde toepassing van PET en PET/CT in de dagelijkse praktijk, en voor enkele ziektebeelden in het bijzonder.

Diagnostiek en behandeling van kanker in het hoofd-hals gebied

Voor het afbeelden van kanker in het hoofd-hals gebied werd de beeldkwaliteit van FDG-PET geoptimaliseerd, en dit leidde tot verbeterde stadiëring en behandeling van patiënten met deze ziekte (hoofdstuk 3). Aangezien deze aanpassingen specifiek zijn voor een bepaalde scanner, werd aanbevolen de optimalisatie uit toe voeren voor alle PET-scanners die voor dit doel worden gebruikt. Helaas werd vastgesteld dat, ondanks deze verbeteringen, beeldvorming met FDG-PET de meer ingrijpende operatieve lymfeklierdissectie nog niet kan vervangen voor de stadiëring van patienten met kanker in het hoofd-hals gebied (hoofdstuk 4).

Vervolgens werd het gebruik van gecombineerde PET- en CT-scans voor het hoofd-hals gebied geoptimaliseerd. Na correctie van een onverwacht verschil in beeldgrootte tussen PET en CT (hoofdstuk 5), en na selectie van de meest geschikte methode om de beelden anatomisch gelijk te positioneren, werd de onnauwkeurigheid in de beeldfusie teruggebracht tot slechts 3 mm. Daarmee werd het combineren van PET en CT gevalideerd voor toepassing bij geavanceerde 3-dimensionale bestraling van kanker in het hoofd-hals gebied (hoofdstuk 6).

Tevens werd geanalyseerd in hoeverre het gebruik van functionele informatie van FDG-PET-beelden invloed had op het intekenen van bestralingsvelden. Daarbij bleek dat de impact van de PET-beelden potentieel groot was, maar dat verschillende methoden om de grenzen van ziektelokalisaties aan te geven onderling sterk kunnen verschillen (hoofdstuk 7). Daarom werd geconcludeerd dat het gebruik van PET-beelden voor planning van bestraling in het hoofd-hals gebied nader onderzoek behoeft, en nog niet als gevalideerde techniek kan worden beschouwd.

Het detecteren van lymfekliermetastasen door het afbeelden van metabolisme met FDG-PET kan worden vertroebeld door de aanwezigheid van ontstekingverschijnselen in deze lymfeklieren. Daarom werd een radioactieve stof die alleen in delende cellen wordt opgenomen, Fluor-18 gekoppeld aan thymidine (FLT), onderzocht als alternatieve tracer. Helaas bleek daarbij dat ook FLT fout-positieve signalen kan geven, doordat het ook wordt opgenomen in zich delende afweercellen in lymfeklieren (hoofdstuk 8). Voor klierstadiëring bij patiënten met kanker in het hoofd-hals gebied kan daarom FDG-PET niet vervangen worden door FLT-PET. Naar de ideale PET-tracer voor deze specifieke toepassing moet nog verder worden gezocht.

Diagnostiek en behandeling van levermetastasen bij darmkanker

De combinatie van PET- en CT-beelden wordt steeds vaker gebruikt voor diagnostiek van patiënten met kanker van de dikke darm, met name als deze ziekte terugkeert na een eerdere behandeling. Deze aanpak kent echter ook enkele nadelen, met name in het gebied van de bovenbuik, waar adembewegingen verschillen tussen de beelden kunnen veroorzaken. In hoofdstuk 9 werd het gebruik van zowel afzonderlijke als gecombineerde PET/CT-beelden voor deze toepassing uitgebreid besproken.

De combinatie van PET- en CT-beelden van afzonderlijke scanners werd geoptimaliseerd voor het afbeelden van de lever en eventuele metastasen van darmkanker daarin. Hiervoor werd een orgaanspecifieke methode voor het anatomisch gelijk positioneren van de beelden ontwikkeld (hoofdstuk 10), met een onnauwkeurigheid van slechts enkele millimeters. Daarmee werd deze aanpak gevalideerd voor klinische toepassing, hoewel de procedure arbeidsintensief en gecompliceerd bleek ten opzichte van het gebruik van een gecombineerde PET/CT scanner.

Vervolgens werd de performance van een gecombineerde PET/CT-scanner voor deze vraagstelling onderzocht (hoofdstuk 11). Daarbij bleek dat anatomische verschillen tussen de beelden ernstig konden zijn, en dat deze veroorzaakt werden door ademhalingsverschillen. Klinisch relevante verschillen traden op in 40-55 procent van de gevallen, afhankelijk van het soort ademhalings instructie dat tijdens het scannen aan de patient werd gegeven. Als gevolg daarvan traden ook potentieel ernstige verminkingen op in de uiteindelijke PET- en CT-beelden, die alleen omzeild konden worden door het beoordelen van PET-beelden die niet waren gecorrigeerd voor verzwakking van radioactieve straling. Op basis van deze bevindingen werden aanbevelingen geformuleerd voor het optimaal gebruiken van een gecombineerde PET/CT scanner, en het beoordelen van de beelden daarvan.

Ten slotte werd het meten van afwijkingen binnen de lever geoptimaliseerd. Een nieuwe methode, specifiek voor het afgrenzen en meten van metastasen van darmkanker binnen de lever, werd ontworpen (hoofdstuk 12). In tegenstelling tot bestaande technieken bleek deze methode bij fantoomproeven geschikt om levermetastasen te evalueren onafhankelijk van achtergrondsignaal, het formaat van de afwijkingen en de intensiteit van de afwijkingen. Vervolgens werd de methode met succes toegepast op FDG-PET-scans van patiënten, in een vergelijking met bestaande technieken. Door controle met bevindingen bij operatie kon de nauwkeurigheid van het meten van het formaat van levermetastasen op PET-beelden met deze methode worden gevalideerd (hoofdstuk 13).

16.2 Conclusies

Verschillende verbeteringen in het gebruik van PET en gecombineerde PET/CT werden gepresenteerd in dit proefschrift. Het verkrijgen, bewerken, beoordelen en meten van beelden werd geoptimaliseerd en gevalideerd voor verschillende toepassingen, met name op het gebied van kanker in het hoofd-hals gebied en levermetastasen van darmkanker.

Echter, sommige kwesties blijven nog onduidelijk of onopgelost. De beeldresolutie van functionele beeldvorming met PET blijft relatief beperkt in vergelijking met CT en MRI. Geïntegreerde beeldvorming met PET/CT blijft een risico houden op artefacten en anatomische verschillen, ondanks de beschreven optimalisaties. Het lokaliseren en afgrenzen van afwijkingen op PET-beelden blijft een goede benadering, ongeacht de uitvoering met separate of geïntegreerde scanners.

Bij de introductie van nieuwe technieken in de medisch diagnostiek blijft kennis van de voor- en nadelen de beste garantie voor adequaat gebruik en goede interpretatie van beelden. Verder onderzoek zal noodzakelijk blijven, evenals het ontwikkelen van protocollen die regionale verschillen verminderen en grootschalig vergelijkbaar onderzoek mogelijk maken, in de huidige tijd van stormachtige technische ontwikkelingen.

Dankwoord

Promoveren is in de eerste plaats erkennen dat je niets weet. Dan kennis vergaren, daarmee wetenschap bedrijven, ergens halverwege het traject een diep dal doorkomen, om uiteindelijk het geheel te overzien. Zo verging het mij ook, en in iedere fase hebben velen me geholpen, waarvoor mijn dank. Een aantal van hen wil ik graag benoemen.

Beste Wim, professor Oyen, jij weet dat er voor een promovendus geen grotere waarden zijn dan plezier en vertrouwen in het onderzoek. Jouw begeleiding, sturend en kritisch waar nodig, maar vooral ook vrij waar mogelijk, heeft de afronding van mijn promotie aangenaam én mogelijk gemaakt.
Beste Frans, professor Corstens, ik ben je zeer erkentelijk voor alle geboden kansen: de opleiding, het eerste artikel, de eerste internationale wetenschappelijke voordracht, het internationaal doceren.
Beste Jorn, Dr. van Dalen, wij deelden onze interesse voor beeldfusie, jij vanuit het fysisch perspectief en ik vanuit het klinisch perspectief, en ik geloof dat dit synergistisch heeft gewerkt. Bovendien heeft elke promovendus een maatje nodig, of meer nog, een voorbeeld.

Net als de promovendus zelf, is zijn afdeling onderdeel van een groter geheel. Beeldvormend onderzoek overschrijdt afdelingsgrenzen, support van zusterdisciplines is onontbeerlijk. De voor mijn onderzoek belangrijkste afdelingen wil ik danken.
De afdeling radiotherapie is historisch gezien, en in het licht van de huidige ontwikkelingen nu opnieuw, een echte geestverwant van de nucleaire geneeskunde. Ons gemeenschappelijke belang in non-invasieve weefselkarakterisering heeft vruchten afgeworpen, en niet in de laatste plaats door de prettige samenwerking. Beste Hans, prof. dr. Kaanders, dank voor je objectieve visie op mijn werk, vanuit de radiotherapie. Beste Dominic, Esther en Aswin, ik ben vereerd dat ik het onderzoek, zowel de successen als de problemen, met jullie mocht delen.
De afdeling radiologie is eveneens een echte zusterdiscipline. De toekomst kan niet anders dan verregaande samenwerking brengen. Beste Nico en HenkJan, als grondleggers van de beeldfusie in het UMC St Radboud was jullie bijdrage onmisbaar, zeker in de beginfase van mijn onderzoek. Daarnaast dank ik mede-auteurs radiologen prof. dr. Barentsz en dr. Jager, voor hun gezond kritische opstelling naar de nucleaire geneeskunde.
Goed wetenschappelijk onderzoek is onlosmakelijk verbonden met goede patiëntenzorg, waar de belangrijke vragen worden opgeworpen en hypothesen kunnen worden getoetst. In dat kader wil ik de afdelingen keel-, neus-, en oorheelkunde (speciaal Bart Wensing, prof. dr. van den Hoogen, prof. dr. Marres), en de afdeling oncologische heelkunde (speciaal Bas Wiering en dr. Theo Ruers) lof toezwaaien.

De afdeling Nucleaire geneeskunde van het UMC St Radboud – mijn afdeling – verdient speciale aandacht. Gewaardeerd binnen de eigen kliniek, gewaardeerd in de internationale wetenschap. Waar iedere patiënt wordt ontvangen met een vriendelijk woord, en met verstand van zaken. De samenwerking altijd gezond en plezierig, zowel op de werkvloer als daarbuiten. Beste collega’s, nogmaals dank voor de geweldige jaren.

Mijn promotie was niet mogelijk geweest zonder enig vermogen tot relativeren, werklust, plezier, en soms ook het incasseren van tegenslagen en kritiek. Wat ik daarvan in mij draag dank ik aan mijn ouders, Jeanne en Jacques. De inspiratie en richting erbij dank ik aan mijn vrienden, in het bijzonder Stan. En met jou, liefste Lonneke, kon ik het allemaal delen.

Curriculum Vitae

Wouter Victor Vogel werd geboren op 16 juni 1973 te ‘s-Hertogenbosch. In 1991 behaalde hij het havo-diploma aan het Rodenborch College te Rosmalen, en in 1994 het vwo-diploma aan het Cranenveld College te Nijmegen. In maart 2001 werd het artsexamen afgelegd aan de Katholieke Universiteit Nijmegen.

Na een keuze-coschap en een korte periode AGNIO-schap volgde Wouter van juli 2001 tot juli 2005 de specialisatie nucleaire geneeskunde in het UMC St. Radboud bij prof. dr. F.H.M. Corstens. In het derde jaar van deze opleiding werd gestart met het in dit proefschrift beschreven promotie-onderzoek, bij prof. dr. W.J.G. Oyen, in samenwerking met de afdelingen radiotherapie, keel-, neus- en oorheelkunde, oncologische heelkunde en radiologie. Vanaf het vierde jaar opleiding was Wouter ook actief als docent binnen de geneeskunde opleiding, (na)scholing van medisch nucleair werkers, en het Europese onderwijscentrum voor PET en PET/CT van de Europese vereniging voor nucleaire geneeskunde (EANM) te Wenen.

In Nijmegen was het leven goed, maar Lonneke wees een nieuwe weg. Sinds april 2007 is Wouter werkzaam als nucleair geneeskundige in het Nederlands Kanker Instituut - Antoni van Leeuwenhoek ziekenhuis, te Amsterdam.

List of publications

1. van der Maas AA, Vogel W. Computerized charting of case reports. Med Inform Internet Med. 2000;25(1):45-61.
2. Vogel WV, Oyen WJ, Barentsz JO, Kaanders JH, Corstens FH. PET/CT: panacea, redundancy, or something in between? J Nucl Med. 2004;45 Suppl 1:15S-24S.
3. van Dalen JA, Vogel W, Huisman HJ, Oyen WJ, Jager GJ, Karssemeijer N. Accuracy of rigid CT-FDG-PET image registration of the liver. Phys Med Biol. 2004;49(23):5393-5405.
4. Vogel WV, van Dalen JA, Huisman H, Oyen WJ, Karssemeijer N. Sliced alternating DICOM series: convenient visualisation of image fusion on PACS. Eur J Nucl Med Mol Imaging. 2005;32(2):247-248.
5. Vogel WV, Wensing BM, van Dalen JA, Krabbe PF, van den Hoogen FJ, Oyen WJ. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. Eur J Nucl Med Mol Imaging. 2005;32(11):1276-1282.
6. Vogel WV, Wiering B, Corstens FH, Ruers TJ, Oyen WJ. Colorectal cancer: the role of PET/CT in recurrence. Cancer Imaging. 2005;5 Spec No A:S143-149.
7. Wensing BM, Vogel WV, Marres HA, Merkx MA, Postema EJ, Oyen WJ, van den Hoogen FJ. FDG-PET in the clinically negative neck in oral squamous cell carcinoma. Laryngoscope. 2006;116(5):809-813.
8. Vogel WV, van Dalen JA, Schinagl DA, Kaanders JH, Huisman H, Corstens FH, Oyen WJ. Correction of an image size difference between positron emission tomography (PET) and computed tomography (CT) improves image fusion of dedicated PET and CT. Nucl Med Commun. 2006;27(6):515-519.
9. Secondary parkinsonism in childhood: A rare complication after radiotherapy. Pediatr Neurol. 2006;34(6):495-498.
10. Vogel WV, van Dalen JA, Oyen WJ. Adequate evaluation of image registration in hybrid PET/CT. J Nucl Med. 2006;47(9):1556; author reply 1557-1558.
11. Vogel WV, van Riel PL, Oyen WJ. FDG-PET/CT can visualise the extent of inflammation in rheumatoid arthritis of the tarsus. Eur J Nucl Med Mol Imaging. 2007;34(3):439.
12. Vogel WV, Schinagl DA, Van Dalen JA, Kaanders JH, Oyen WJ. Validated image fusion of dedicated PET and CT for external beam radiation therapy in the head and neck area. Q J Nucl Med Mol Imaging. 2007 Mar 25.
13. van Dalen JA, Visser EP, Vogel WV, Corstens FH, Oyen WJ. Impact of Ge-68/Ga-68-based versus CT-based attenuation correction on PET. Med Phys. 2007;34(3):889-897.

14. van Dalen JA, Hoffmann AL, Dicken V, Vogel WV, Wiering B, Ruers TJ, Karssemeijer N, Oyen WJ. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. Nucl Med Commun. 2007;28(6):485-493.

15. Troost EG, Vogel WV, Merkx MA, Slootweg PJ, Marres HA, Peeters WJ, Bussink J, van der Kogel AJ, Oyen WJ, and Kaanders JH. 18F-FLT PET does not discriminate between reactive and metastatic lymph nodes in primary head and neck cancer patients. J Nucl Med. 2007;48(5):726-735.

16. Vogel WV, van Dalen JA, Wiering B, Huisman H, Corstens FH, Ruers TJ, and Oyen WJ. Evaluation of image registration in PET/CT of the liver and recommendations for optimized imaging. J Nucl Med. 2007;48(6):910-919.

17. van Dalen JA, Vogel WV, Corstens FH, and Oyen WJ. Multi-modality nuclear medicine imaging: artefacts, pitfalls and recommendations. Cancer Imaging. 2007;7:77-83.

Stellingen behorende bij het proefschrift

1. Instituten met een goede CT scanner, een goede PET scanner, en adequate interpretatie van de beelden, zijn niet minderwaardig aan instituten met een geïntegreerde PET/CT scanner. (Dit proefschrift)
2. Het corrigeren van PET-beelden voor fotonverzwakking middels CT-beelden vermindert de beeldkwaliteit, en moet derhalve verworpen worden zodra het economisch voordeel ervan wegvalt. (Dit proefschrift)
3. De toegevoegde waarde van beeldfusie ligt vaak niet in integratie van verschillende soorten beelden, maar in de daaropvolgende communicatie tussen de bijbehorende specialisten. (Dit proefschrift)
4. Door de vele beschikbare diagnostische technieken en strategieën, en de snelle implementatie daarvan, is het moeilijk om gedegen wetenschappelijk onderzoek te verrichten naar de werkelijke waarde ervan. (Dit proefschrift)
5. Van behandelend artsen kan niet langer worden verwacht dat zij te allen tijde weten welk beeldvormend onderzoek, en in welke volgorde, het beste kan worden aangevraagd. (Dit proefschrift)
6. De term ‘nucleaire geneeskunde’ biedt alvast semantische ruimte voor integratie met ‘kernspin resonantie’ en ‘atomaire fluorescentie’.
7. De werving en verdeling van geld voor medisch onderzoek berust thans teveel op de mediageniciteit van de betreffende ziekten.
8. Als uw arts ‘PubMed’ als bookmark heeft, is uw kans op genezing groter.
9. De som der eigenbelangen is niet het algemeen belang.
10. Na zichzelf zal de mens ook het klimaat een schoonheidsideaal opleggen.
11. Het afwijzen van de evolutietheorie biedt kennelijk een evolutionair voordeel.
12. There are in fact two things, science and opinion; the former begets knowledge, the latter ignorance. (Hippocrates, 460 BC - 377 BC)
13. Treasure the love you receive, it will survive long after your good health has vanished. (Augustine Mandino, 1923 - 1996)
14. Wie nimmer het nest verlaat, zal nooit weten of hij kan vliegen.
    (Jeanne en Jacques Vogel, in 1992)

Wouter Vogel, 2007